

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2001, 03:46:36 ; Search time 1381.47 Seconds
(without alignments)
4060.207 Million cell updates/sec

Title: us-09-402-488a-1

Perfect score: 1096
Sequence: 1 ATGTCCCTACTAGCTTA.....ATCCACATCTTAGACGG 1096

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 118133 seqs, 2558675100 residues

Total number of hits satisfying chosen parameters: 2236266

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba1: *
2: gb_ba2: *
3: gb_om: *
4: gb_ov: *
5: gb_ph: *
6: gb_pl1: *
7: gb_pl2: *
8: gb_pl3: *
9: gb_pr2: *
10: gb_pr3: *
11: gb_ro: *
12: gb_sy: *
13: gb_un: *
14: em_fun: *
15: em_hum1: *
16: em_hum2: *
17: em_in: *
18: em_om: *
19: em_or: *
20: em_ov: *
21: em_pat: *
22: em_ph: *
23: em_pl: *
24: em_ro: *
25: em_sts: *
26: em_sy: *
27: em_un: *
28: em_vl: *
29: gb_ba3: *
30: gb_in1: *
31: gb_in2: *
32: gb_in3: *
33: gb_pl3: *
34: gb_pr4: *
35: em_ba1: *
36: em_ba2: *
37: em_hc91: *
38: em_hc92: *
39: em_hc93: *
40: em_hc94: *
41: em_hc95: *
42: em_hc96: *
43: em_hc97: *

44: em_hc98: *
45: em_hc99: *
46: em_hc10: *
47: em_hum3: *
48: em_hum4: *
49: em_hum5: *
50: em_hum6: *
51: gb_pr5: *
52: gb_pr6: *
53: gb_pr7: *
54: gb_hc91: *
55: gb_hc92: *
56: gb_hc93: *
57: gb_hc94: *
58: gb_hc95: *
59: gb_hc96: *
60: gb_hc97: *
61: gb_hc98: *
62: gb_hc99: *
63: gb_hc10: *
64: gb_hc91: *
65: gb_hc92: *
66: gb_hc93: *
67: gb_hc94: *
68: gb_hc95: *
69: gb_hc96: *
70: gb_hc97: *
71: gb_hc98: *
72: gb_hc99: *
73: gb_hc10: *
74: gb_hc91: *
75: gb_hc92: *
76: gb_hc93: *
77: gb_hc94: *
78: gb_hc95: *
79: gb_hc96: *
80: gb_hc97: *
81: gb_hc98: *
82: gb_hc99: *
83: em_hc90: *
84: gb_hc924: *
85: gb_pr8: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query	Length	DB ID	Description
1	708.6	64.7	4968	12 XXU13855	U13855 pGEX-4T3 c1
2	697.6	63.6	4969	12 XXU13853	U13853 pGEX-4T1 c1
3	696.6	63.6	4970	12 XXU13854	U13854 pGEX-4T2 c1
4	685	62.5	4972	12 XXU13856	U13856 pGEX-5X1 c1
5	682.4	62.3	4947	12 XXU13849	U13849 pGEX-1 Lamb
6	681	62.1	3417	12 AB014641	AB014641 Cloning v
7	681	62.1	4948	12 XXU13850	U13850 pGEX-2T c10
8	681	62.1	5042	12 EVU84572	U84572 Expression
9	681	62.1	5052	12 EVU84571	U84571 Expression
10	681	62.1	5501	12 EVU85201	U85201 Expression
11	681	62.1	5504	12 EVU85202	U85202 Expression
12	681	62.1	5516	12 EVU85203	U85203 Expression
13	681	62.1	5738	12 EVU85204	U85204 Expression
14	681	62.1	5756	12 EVU85205	U85205 Expression
15	681	62.1	5990	12 EVU85206	U85206 Expression
16	679.2	62.0	4973	12 XXU13857	U13857 pGEX-5X2 c1
17	678.4	61.9	1281	81 A47307	A47307 Sequence 2
18	678.2	61.9	1740	81 AR067794	AR067794 Sequence
19	678	61.9	9760	81 EVU86615	EVU86615 Expression
20	678	61.9	9763	12 AF126280	AF126280 Expression
21	677.4	61.8	4969	12 XXU13851	U13851 pGEX-2TK c1

22 677 61.8 1119 81 E13745
23 677 61.8 10765 12 EVU67875
24 676.2 61.7 7805 33 YSCCJ2INDU
25 675.6 61.6 4983 12 CVU78874
26 675 61.6 4974 12 XXU13858
27 674.6 61.6 3378 12 AF149265
28 673.8 61.5 3966 12 EVU81274
29 668.2 61.0 4984 12 CVU78872
30 667.4 60.9 897 81 E16870
31 664.2 60.6 897 21 E09612
32 664 60.6 4985 12 CVU78873
33 663.8 60.6 5487 12 PGH23813
34 663.4 60.5 1128 21 E09265
35 663.4 60.5 4952 12 XXU13852
36 662.8 60.5 2095 81 AR026350
37 661.4 60.3 3950 12 EVU81273
38 661 60.3 4932 12 SYNECOPGX
39 660.2 60.2 3970 12 EVU81275
40 658.6 60.1 4905 12 ASPGST
41 655.2 59.8 5037 12 AF097411
42 655 59.8 739 32 SCMAg
43 654 59.7 654 81 AR036684
44 652.2 59.5 897 21 E09610
45 651 59.4 9750 12 AF025397

ALIGNMENTS

RESULT 1

LOCUS XXU13855 4968 bp DNA circular SYN
DEFINITION pGEX-4T3 cloning vector, complete sequence.
ACCESSION U13855
VERSION U13855.1 GI:595729
KEYWORDS glutathione S-transferase; beta-lactamase; lac repressor.
SOURCE unidentified cloning vector.
ORGANISM artificial sequence; vectors.
REFERENCE 1 (bases 1 to 4968)
AUTHORS Malone,J.A.
TITLE pGEX-4T3: A cloning vector for the inducible expression of genes as
glutathione S-transferase fusion proteins containing a thrombin
cleavage site
JOURNAL Unpublished (1994)
REFERENCE 2 (bases 1 to 4968)
AUTHORS Smith,D.B. and Johnson,K.S.
TITLE Single-step purification of polypeptides expressed in Escherichia
coli as fusions with glutathione S-transferase
JOURNAL Gene 67 (1), 31-40 (1988)
MEDLINE 88329742
REFERENCE 3 (bases 264 to 911)
AUTHORS Smith,D.B., Davern,K.M., Board,P.G., Tiu,W.U., Garcia,E.G. and
Mitchell,G.F.
TITLE Mr 26,000 antigen of Schistosoma japonicum recognized by resistant
WEHI 129/J mice is a parasite glutathione S-transferase [published
erratum appears in Proc Natl Acad Sci U S A 1987 Sep;84(18):6541]
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 83 (22), 8703-8707 (1986)
MEDLINE 87041520
REFERENCE 4 (bases 881 to 911)
AUTHORS Smith,D.B., Davern,K.M., Board,P.G., Tiu,W.U., Garcia,E.G. and
Mitchell,G.F.
TITLE Correction: Mr 26,000 antigen of Schistosoma japonicum recognized
by resistant WEHI 129/J mice is a parasite glutathione
S-transferase
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 84, 6541-6541 (1987)
REFERENCE 5 (bases 1 to 4968)
AUTHORS Malone,J.A.
TITLE Direct Submission
JOURNAL Submitted (13-AUG-1994) James A. Malone, International Technical
Services, Molecular Biology Reagents Division, Pharmacia Biotech
Inc., 2202 N. Bartlett Ave., Milwaukee, WI 53202-1009, USA
FEATURES Location/Qualifiers

source 1..4968
/organism="unidentified cloning vector"
/db_xref="taxon:45196"
/lab_host="Escherichia coli"
183..211
/standard_name="lac"
/note="lac promoter for inducible expression of
glutathione S-transferase"
217..237
protein_bind
CDS
258..980
/bound_molec="lac repressor protein"
/citation=[3]
/citation=[4]
/codon_start=1
/transl_table=11
/product="glutathione S-transferase"
/protein_id="AA57104.1"
/db_xref="GI:595730"
/translation="MSPILGKWKIKGLVOPRLLELYEERKEEHLERDEGDKRNRK
KFELEPEPNLPYIDGDKVLTOSMAIIRYIADKHNMLGCPKRRARISMLEGAVDI
RYGVSRIASKEDETLAKVDPLSKPEAKKEDEKCHRTYINQDVTHTPDMYDALD
VLIYMDPACLDAPFKYCFKRIKRIATPQIDKYSKYIAMPLOGWQATRGCGDHPK
SDLPKGSPPNSRVDSGRIVYD"
918..935
misc_feature
/note="encodes thrombin recognition site"
930..965
misc_feature
/note="Multiple Cloning Site (MCS); contains the unique
restriction sites BamHI, EcoRI, SmaI, SalI, XhoI, and
NotI"
1306..2236
/gene="bla"
/gene="bla"
/gene="bla"
1376..2236
/gene="bla"
/transl_table=1
/codon_start=1
/product="beta-lactamase"
/protein_id="AA57105.1"
/db_xref="GI:595731"
/translation="MSTQHRVALIPFAECPLVPAPHTLVKVKDAEDOLGARVY
IEIDLNSGKLTIESRPERPMPSTFVLLGAVLSVDAGOEOLGRIRIYSONDLVE
YSPVEKHLIDGMVRELCSAITSNTANLLTIGPKELTARLHMGHVPRL
DRPEPELNEAIPDERDRTTMAPMATLTKLLIGELTTLAROOIIDMADYVAGL
LRSAIPAGWFIADKSGGERSGCIITALADPDKPSRIYVITTTGSATYDENRQTA
ETGASLIRKH"
2994
rep_origin
/note="base 2994 represents the first base of the newly
synthesized single strand"
3317..4399
/direction=right
3317..4399
/gene="lacIq"
3317..4399
/gene="lacIq"
/transl_table=11
/codon_start=1
/product="lac repressor"
/protein_id="AA57106.1"
/db_xref="GI:595732"
/translation="MKPVTLYDAEYAGVSYQTVSRVNOASHVSAKTRKVEAAMAE
LNYIPNVAOOLAGKOSLIGVATSSALAPSOIYAATIRADOLGASVVSVMERS
GVACAAAVNLLAORVSGLIINYPIDDDAIVAEACTNVPALFLFDYSQPTINSIT
FSHEDGTRICVEHLVALGHOAILLACGPPSSVSARILRAGWHTYLTIRNOIPIAEBEG
DMSAGSGFOOTKMDNGIYPTAMLVANDMAIGARITTESGLRVGADISVGYDPT
EDSSCYIPLTTIKQDFRLIGQTSVDRLQLSOGQAVKGNQLPVSILVYKRTYLAANT
QRTSPALADSLMQLAROVSRLESQ"
1202 c 1291 g 1250 t

BASE COUNT

ORIGIN 1225 a 1291 g 1250 t

Query Match 64.7%; Score 708.6; DB 12; Length 4968;
Best Local Similarity 99.4%; Pred. No. 3.7e-172;
Matches 711; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY	1	ATGCCCTTACAGTAGTTATTGGAAATTAAGGCGCTTGTCACCCACCACGTGCATTC	60
Db	258	ATGTCOCCTATACHTAAGTATTGGAAAAATTAAAGGCGCTTGTCACCCACCACGTGCATTC	317
OY	61	TTCGAATAATCTTGAAGAAAAATATGAAGAGCATTTGTATGAGCCGATGAAGGTATAA	120
Db	318	TTTGCAATATCTTGAAGAAMAAATATGAAGAGCATTTGTATGAGCCGATGAAGGTATAA	377
OY	121	TGGCGAACAACAAAGTTTGAATTTGGGTTTGGAGTTTCCCATCTTCCTTATTATTGAT	180
Db	378	TGGCGAACAACAAAGTTTGAATTTGGGTTTGGAGTTTCCCATCTTCCTTATTATTGAT	437
OY	181	GCTGATGTTAAATTACACAGCTCTATGGCCATCATACGTTATATAGCTGACAGACAAC	240
Db	438	GCTGATGTTAAATTATACACAGCTCTATGGCCATCATACGTTATATAGCTGACAGACAAC	497
OY	241	ATGTTGGGTGCTGTCTCACAAGAGCGCTCAGAGATTTCATATGCTTGAAGAGAGGGTTTTG	300
Db	498	ATGTTGGGTGCTGTCTCACAAGAGCGCTCAGAGATTTCATATGCTTGAAGAGAGGGTTTTG	557
OY	301	GATATTACATACGGGTTCGAGATTTCGATTAATAGTAAAGACTTTGAAACTCTCAAAGTT	360
Db	558	GATATTACATACGGGTTCGAGATTTCGATTAATAGTAAAGACTTTGAAACTCTCAAAGTT	617
OY	361	GATTTTCTTACCAAGCTACCTGAAAATGCTGAAAATGTTTCGAAGATCGTTTATGTCATAA	420
Db	618	GATTTTCTTACCAAGCTACCTGAAAATGCTGAAAATGTTTCGAAGATCGTTTATGTCATAA	677
OY	421	ACATATATTTAAATGCTGATCATGTATAACCCATCCCTGCATTCATGTTTATGACCGCTTGAT	480
Db	678	ACATATATTTAAATGCTGATCATGTATAACCCATCCCTGCATTCATGTTTATGACCGCTTGAT	737
OY	481	GTTGTTTATACATRGACCAATGTCCTGATGAGTGGTTTCCCAAATTTAGTTGTTTAA	540
Db	738	GTTGTTTATACATRGACCAATGTCCTGATGAGTGGTTTCCCAAATTTAGTTGTTTAA	797
OY	541	AAACGATTTGAAGCTATCCCAACAATTTGATTAAGTACTTGAAATCCAGCAAGATATTAGA	600
Db	798	AAACGATTTGAAGCTATCCCAACAATTTGATTAAGTACTTGAAATCCAGCAAGATATTAGA	857
OY	601	TGCGCTTTGACAGGCGTGCAAGCACGTTTGGTGTGGCGACCATCTCCAAATTCGGAT	660
Db	858	TGCGCTTTGACAGGCGTGCAAGCACGTTTGGTGTGGCGACCATCTCCAAATTCGGAT	917
OY	661	CTGCTTCGCGGTGATATCCCGCAATTTCCCGGGTCGACCTGAGCGCGCGCTAGA	715
Db	918	CTGCTTCGCGGTGATATCCCGCAATTTCCCGGGTCGACCTGAGCGCGCGCATCGTGA	972
RESULT 2			
XXU13853			
LOCUS	XXU13853	4969 bp	DNA circular SYN 13-DEC-1994
DEFINITION	pSEX-4T1 cloning vector, complete sequence.		
VERSION	U13853		
KEYWORDS	U13853.1 GI:595721		
SOURCE	glutathione S-transferase; beta-lactamase; lac repressor.		
ORGANISM	unidentified cloning vector.		
REFERENCE	unidentified cloning vector.		
AUTHORS	artificial sequence; vectors.		
TITLE	1 (bases 1 to 4969)		
	Malone,J.A.		
	pEX-4T1: A cloning vector for the inducible expression of genes as		
	glutathione S-transferase fusion proteins containing a thrombin		
	cleavage site		
JOURNAL	Unpublished (1994)		
REFERENCE	2 (bases 1 to 4969)		
AUTHORS	Smith,D.B. and Johnson,K.S.		
TITLE	Single-step purification of polypeptides expressed in Escherichia		
	coli as fusions with glutathione S-transferase		
JOURNAL	Gene 67 (1), 31-40 (1988)		
MEDLINE	88329742		
REFERENCE	3 (bases 264 to 911)		

```

AUTHORS Smith,D.B., Daveern,K.M., Board,P.G., Tiu,W.U., Garcia,E.G. and Mitchell,G.F.
TITLE ME 26,000 antigen of Schistosoma japonicum recognized by resistant WHI 129/J mice is a parasite glutathione S-transferase [published erratum appears in Proc Natl Acad Sci U S A 1987 Sep;84(18):6541]
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 83 (22), 8703-8707 (1986)
MEDLINE 87041520
REFERENCE 4 (bases 881 to 911)
AUTHORS Smith,D.B., Daveern,K.M., Board,P.G., Tiu,W.U., Garcia,E.G. and Mitchell,G.F.
TITLE Correction: Mr 26,000 antigen of Schistosoma japonicum recognized by resistant WHI 129/J mice is a parasite glutathione S-transferase
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 84, 6541-6541 (1987)
REFERENCE 5 (bases 1 to 4969)
AUTHORS Malone,J.A.
TITLE Direct Submission
JOURNAL Submitted (19-AUG-1994) James A. Malone, International Technical Services, Molecular Biology Reagents Division, Pharmacia Biotech Inc., 2202 N. Bartlett Ave., Milwaukee, WI 53202-1009, USA
FEATURES
source 1..4969
/organism="unidentified cloning vector"
/db_xref="taxon:45196"
/lab_host="Escherichia coli"
183..211
/standard_name="lac"
/promoter="/note="tac promoter for inducible expression of glutathione S-transferase"
217..237
/bound_moiety="lac repressor protein"
258..977
/citation=[3]
/citation=[4]
/codon_start=1
/transl_table=11
/product="glutathione S-transferase"
/protein_id="AA57098.1"
/db_xref="GI:595722"
/translacion="MSPILGVWKIKGLVQPTRLLEYLEEKYEELVERDEGDKRMKNK KFELEGFNLPYVIGDVKLTQSMALIRIYADKHNMLGGCKPRRAETSMLEGVLDIT RGVSRVAYSKDEPTLKYVDFLSKPELKNKPEDELCRTYLLGNCHVTHPPDMLYDALD VLVYMDPMCLDAEPKLYCFKRRIRIADIPQIDKYLKSKYIAMPLOGWATPEGGDPKPS SDIVPQSGPEPFKRLERPHRD"
918..935
/note="encodes {thrombin recognition site"
930..966
/note="Multiple Cloning Site (MCS); contains the unique restriction sites BamHI, EcoRI, SmaI, SalI, XhoI, and NotI; the EcoRI site of pGEX-4T1 is in frame with the EcoRI site of lambda gt11"
1307..2237
/gene="bla"
1307..1335
/gene="bla"
1377..2237
/gene="bla"
1377..2237
/gene="bla"
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="AA57099.1"
/db_xref="GI:595723"
/translacion="MSIOHFRVALIPFEAFCLPFAHPETIVKXDAEDDLGARVGG IELDINGKILLESFREPPEPMSTFVLICGAVALSDAGGBOIGRIHNSDILVEB YSPVTEKHLDGMYRELCSSAATMTSGNTAGNALLITIGKPELTARLHNGDHVTL DKNPELENAIPNDRERTTPAAMATTLTKRLILGELLTLASRODLIMWEDKVGAPD LKRALPGWFIADKSGAGERGSGRIIATLDPGKPSRIWIVYTTGSQATDERNRQIA ELGASLTKH"
2995
/note="base 2995 represents the first base of the newly synthesized single strand"
/direction=right
3318..4400
gene

```


	Query Match	Best local similarity	Matches 711: Conservative	Score 696.6	DB 12	Length 4970
QY	1	ATGCCCCCTAATCTAGGTTATTTGGAATAATTAAGGCGCTTGTGCACACCCATCGACTTCTT	63.6%	99.2%	Pred. No. 4.6e-169	Matches 4: Indels 2: Gaps 1:
Db	258	ATGTCGCCCTTACAGTGTATTTGGAATAATTAAGGCGCTTGTGCACACCCATCGACTTCTT	0	Mismatches 4		
QY	61	TTTGAATATCTTTGAAGAAAATATGAAGACATTTGTATGAGCGCGATGAAGTGATAA				
Db	318	TTTGAATATCTTTGAAGAAAATATGAAGACATTTGTATGAGCGCGATGAAGTGATAA				
QY	121	TGGCGAACAATAAGTTGGAATTTGGGTTTGGAGTTTCCCAATCTTCTTATATATGAT				
Db	378	TGGCGAACAATAAGTTGGAATTTGGGTTTGGAGTTTCCCAATCTTCTTATATATGAT				
QY	181	GGTGATTTAAATTTAACACAGCTCTATGGCCATATACGTTATATAGCTGACACAAGACAAC				
Db	438	GGTGATTTAAATTTAACACAGCTCTATGGCCATATACGTTATATAGCTGACACAAGACAAC				
QY	241	ATGTGGGCGTGTGTCTCCAAAAGAGCGTCCAGAGATTTCAATGCTTTGAAGGACCGCTTTTG				
Db	498	ATGTGGGCGTGTGTCTCCAAAAGAGCGTCCAGAGATTTCAATGCTTTGAAGGACCGCTTTTG				
QY	301	GATATTAATATACGGTGTTCGAGAAATTCGATATAGTAAGAAGCTTTGAANAAGCTTCAAAGTT				
Db	558	GATATTAATATACGGTGTTCGAGAAATTCGATATAGTAAGAAGCTTTGAANAAGCTTCAAAGTT				
QY	361	GATTTCTTAGCAACCTCTGAATTCGTAAGAAATGTTGCGAAGATCGTTATAGTATATAA				
Db	618	GATTTCTTAGCAACCTCTGAATTCGTAAGAAATGTTGCGAAGATCGTTATAGTATATAA				
QY	421	ACATATTTAAATGATGATCATGTAAACCATCTCGACTTCATGTTGTATGACGCTTGTAT				

DB	678	ACATATTAAATGGGATCATATGTAACCATCTGACTCTCATCTGTTATGACCGCTTGAT	737
OY	481	GTGTGTTTATACATGAGACCAATGTGCTCGATGCGGTTCCCAAAATAGTTGTTTAAA	540
DB	738	GTTGTTTATATACATGAGACCAATGTGCTCGATGCGGTTCCCAAAATAGTTGTTTAAA	797
OY	541	AAACGATTTGAAGCATATCCACAAATTTGATAGTACTTGAAATCCAGCAAGTATATACA	600
DB	798	AAACGATTTGAAGCATATCCACAAATTTGATAGTACTTGAAATCCAGCAAGTATATACA	857
OY	601	TGCGCTTTGACAGCGCTGGCAAGCCAGTGTGTTGTTGGCGGACCATTCGCAAAATCGGAT	660
DB	858	TGCGCTTTGACAGCGCTGGCAAGCCAGTGTGTTGTTGGCGGACCATTCCTCCAAATTCGGAT	917
OY	661	CTGTGTTCCGCGGTGATGCC--GAATTCCTCGGCTGACCTGACGCGCGCGCTGAGA	715
DB	918	CTGTGTTCCGCGGTGATGCCAGAGATTTCCCGGCTGACTGACGCGCGCGCATCGTGA	974
RESULT	4		
XXU13856			
LOCUS	XXU13856	4972 bp	DNA circular SYN 13-DEC-1994
DEFINITION	pgEX-5X1	cloning vector,	complete sequence.
ACCESSION	U13856		
VERSION	U13856.1	GI:595733	
KEYWORDS	glutathione S-transferase; beta-lactamase; lac repressor.		
SOURCE	unidentified cloning vector.		
ORGANISM	unidentified cloning vector.		
REFERENCE	1 (bases 1 to 4972)		
AUTHORS	Malone,J.A.		
TITLE	pgEX-5X1: A cloning vector for the inducible expression of genes as		
JOURNAL	glutathione S-transferase fusion proteins containing a factor Xa		
REFERENCE	2 (bases 1 to 4972)		
AUTHORS	Smith,D.B. and Johnson,K.S.		
TITLE	Single-step purification of polypeptides expressed in Escherichia		
JOURNAL	coli as fusions with glutathione S-transferase		
MEDLINE	Gene 67 (1), 31-40 (1988)		
REFERENCE	88329742		
AUTHORS	3 (bases 264 to 911)		
TITLE	Smith,D.B., Davern,K.M., Board,P.G., Tiu,W.U., Garcia,E.G. and		
JOURNAL	Mitchell,G.F.		
MEDLINE	Mr 26,000 antigen of Schistosoma japonicum recognized by resistant		
REFERENCE	WEHI 129/J mice is a parasite glutathione S-transferase [published		
AUTHORS	erratum appears in Proc Natl Acad Sci U S A 1987 Sep;84(18):6541]		
TITLE	Proc. Natl. Acad. Sci. U.S.A. 83 (22), 8703-8707 (1986)		
JOURNAL	87041520		
MEDLINE	4 (bases 881 to 911)		
REFERENCE	Smith,D.B., Davern,K.M., Board,P.G., Tiu,W.U., Garcia,E.G. and		
AUTHORS	Mitchell,G.F.		
TITLE	Correction: Mr 26,000 antigen of Schistosoma japonicum recognized		
JOURNAL	by resistant WEHI 129/J mice is a parasite glutathione		
MEDLINE	S-transferase		
REFERENCE	5 (bases 1 to 4972)		
AUTHORS	Proc. Natl. Acad. Sci. U.S.A. 84, 6541-6541 (1987)		
TITLE	5 (bases 1 to 4972)		
JOURNAL	Malone,J.A.		
MEDLINE	Direct Submission		
REFERENCE	Submitted (13-AUG-1994) James A. Malone, International Technical		
AUTHORS	Services, Molecular Biology Reagents Division, Pharmacia Biotech		
TITLE	Inc., 2202 N. Bartlett Ave., Milwaukee, WI 53202-1009, USA		
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			

```
CDS
/bound-moety="lac repressor protein"
258..980
/citation={3}
/citation={4}
/codon_start=1
/transl_table=1
/product="glutathione S-transferase"
/protein_id="AA57107.1"
/db_xref="GI:595734"
/translacion="MSPILGVYKIKGVOPRLLILEYEEHLYERDGDKRRK
KEELGFEFNPPIYIDGDKITQSMATIRIADKHNMGCPKRAELSMELGVLDI
RYGVSRIAYSKDEFKIVDLKIPKEMKMEDELCHKRYLNGVHPDMLDALD
VVLVMDPCLADAPFKIVCFKRIEAIPOIDKYLKSKSYIAMPLOGWATGEGDHPK
SDLEGRGIDPEFGRLEPRHD"
921..932
/misc_feature
/note="encodes factor Xa recognition site"
934..969
/note="Multiple Cloning Site (MCS): contains the unique
restriction sites BamHI, EcoRI, SmaI, SalI, XhoI, and
NotI; The EcoRI site of pGEX-5x1 is in frame with the
EcoRI site of lambda gtl1"
1310..2240
/gene="bla"
1310..1338
/gene="bla"
1380..2240
/gene="bla"
/codon_start=1
/transl_table=1
/product="beta-lactamase"
/protein_id="AA57108.1"
/db_xref="GI:595735"
/translacion="MSIOHFRVALIPFEAFLPVFAHPETLVKDAEDLGARVY
IELDLSGKLTLESFEPERPEMSTFVLLCGAVLSVDAEQDLGRIRIYSDNLYE
KSPYKEKIDGDKVRELCSAATMSNTANILITIGKEKELTALHNHGDVRL
DRWPELNEALPDEROTTPMAAATTLRLKLTLELLTASRQDLDMEDKAGPL
LRSLPAGWFLADKSGAGERSGRIIAGLGPDRIVITVTGSOATMDERNRQIA
EIGASLIRKHW"
2998
/rep_origin
/note="base 2998 represents the first base of the newly
synthesized single strand"
/direction=right
3321..4403
/gene="lacIq"
3321..4403
/gene="lacIq"
/codon_start=1
/transl_table=1
/product="lac repressor"
/protein_id="AA57109.1"
/db_xref="GI:595736"
/translacion="MKPVTLIDVAEYAGVSYOTVSRVYVNOASHVSAKTRKEVAAAE
LNIIPNVAAQOLAKGOSLITGVATSSIALINPSIYVAIKRAQLGASVYNSVENS
GVEACKAAVHNLAAQVSGILINPLDDODAIAEACTNVPALPLVSDQPLNSII
FSHDDGRIAGEHLVALGHOQIALAGPLSSVSAIRIAGMKYLTNRQIOPAREG
DWSAMSGFOQDMQINIEGIYPTAMLVANDQALGAMRAITSGRGVADISVYDPT
EDSSCIYPLTTIKODPRLLGOTVDRILQISQGAQVKNQLLPVSILKRRKTLAPNT
QTASPRALASIMOLAROVSRLESQ"
BASE COUNT      1228 a      1201 c      1293 g      1250 t
ORIGIN
```

```
Query Match      62.5%: Score 685; DB 12; Length 4972;
Best Local Similarity 98.1%; Pred. No. 4.6e-166;
Matches 705; Conservative 0; Mismatches 10; Indels 4; Gaps 1;
```

```
QY 1 ATGCCCCCTACTAGCTTATTTGAAGGAGCCCTTGCAACCACTGACTCTT 60
|||||
Db 258 ATGCCCCCTACTAGCTTATTTGAAGGAGCCCTTGCAACCACTGACTCTT 317
|||||
QY 61 TTGAATATCTTGAAGAAATATGAAGCATTTGTATGAGCGGATGAAGTGATAAA 120
|||||
Db 318 TTGAATATCTTGAAGAAATATGAAGCATTTGTATGAGCGGATGAAGTGATAAA 377
|||||
```

```
QY 121 TGGCAAAACAAAAGTTGAATGGGTTTGAGATTGCCAATCTCTTATTATTTGAT 180
|||||
Db 378 TGGCAAAACAAAAGTTGAATGGGTTTGAGATTGCCAATCTCTTATTATTTGAT 437
|||||
QY 181 GGTGATGTTAAATTACACAGCTATGSCCATCATAGCTTATATAGTGACACACAAAC 240
|||||
Db 438 GGTGATGTTAAATTACACAGCTATGSCCATCATAGCTTATATAGTGACACACAAAC 497
|||||
QY 241 ATGTGGGTGGTGTGCCAAAAGAGCGTGACAGATTGTAATGCTTGAAGAGCGGTTTG 300
|||||
Db 498 ATGTGGGTGGTGTGCCAAAAGAGCGTGACAGATTGTAATGCTTGAAGAGCGGTTTG 557
|||||
QY 301 GATATTAGATACGCGTGTTCAGAAATTCATATAGTAAGAATTGAAACTCTCAAGTT 360
|||||
Db 558 GATATTAGATACGCGTGTTCAGAAATTCATATAGTAAGAATTGAAACTCTCAAGTT 617
|||||
QY 361 GATTTTCTTAGCAAGCTACCTGAATGCTGAATAATGTTGGAAGATCGTTATGTCATAA 420
|||||
Db 618 GATTTTCTTAGCAAGCTACCTGAATGCTGAATAATGTTGGAAGATCGTTATGTCATAA 677
|||||
QY 421 ACATTTTAATAGTGATCATGTATACCATGCTGACTTCATAGTTATGACGCTTGAT 480
|||||
Db 678 ACATTTTAATAGTGATCATGTATACCATGCTGACTTCATAGTTATGACGCTTGAT 737
|||||
QY 481 GTTGTTTATACATGAGACCAATGCGCTGAGATGCTTCCCAAAATAGTTGTTTAAA 540
|||||
Db 738 GTTGTTTATACATGAGACCAATGCGCTGAGATGCTTCCCAAAATAGTTGTTTAAA 797
|||||
QY 541 AAACGTAATGAGCTATATCCACAAATTTATAGTCTGAAATCCAGCAAGTATATACCA 600
|||||
Db 798 AAACGTAATGAGCTATATCCACAAATTTATAGTCTGAAATCCAGCAAGTATATACCA 857
|||||
QY 601 TGGCCTTGCAAGGCGTGCAAGCCAGCTGTGTGTGTGGGACCATCTCCAAAATCGGAT 660
|||||
Db 858 TGGCCTTGCAAGGCGTGCAAGCCAGCTGTGTGTGTGGGACCATCTCCAAAATCGGAT 917
|||||
QY 661 CTGTTCCCGCGT---GGATCCCGCAATTCGCGGCTGACATCGAGCGGCGCGTGTAGA 715
|||||
Db 918 CTGATCGAAGGTCGTGGGATCCCGCAATTCGCGGCTGACATCGAGCGGCGCGTGTAGA 976
|||||
```

RESULT 5
XXU13849
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
REFERENCE

XXU13849 4947 bp DNA circular SYN 13-DEC-1994
pGEX-1 Lambda T cloning vector, complete sequence.
U13849
U13849.1 GI:595705
glutathione S-transferase; beta-lactamase; lac repressor.
unidentified cloning vector.
unidentified cloning vector.
artificial sequence; vectors.
1 (bases 1 to 4947)
Malone,J.A.
pGEX-1 Lambda T: A cloning vector for the inducible expression of
genes in glutathione S-transferase fusion proteins containing a
thrombin cleavage site
Unpublished (1994)
2 (bases 1 to 4947)
Smith,D.B., and Johnson,K.S.
Single-step purification of polypeptides expressed in Escherichia
coli as fusions with glutathione S-transferase
Gene 67 (1), 31-40 (1988)
88329742
3 (bases 264 to 911)
Smith,D.B., Davenport,K.M., Board,P.G., Tlu,W.U., Garcia,E.G. and
Mitchell,G.F.
Mr 26,000 antigen of Schistosoma japonicum recognized by resistant
WEHI 129/J mice is a parasite glutathione S-transferase [published
erratum appears in Proc Natl Acad Sci U S A 1987 Sep;84(18):6541]
JOURNAL
MEDLINE
REFERENCE
Proc. Natl. Acad. Sci. U.S.A. 83 (22), 8703-8707 (1986)
87041520
4 (bases 881 to 911)

AUTHORS Smith,D.B., Davern,K.M., Board,P.G., Tin,W.U., Garcia,E.G. and Mitchell,G.F.
TITLE Correction: Mr 26,000 antigen of Schistosoma japonicum recognized by resistant WEHI 129/J mice is a parasite glutathione S-transferase
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 84, 6541-6541 (1987)
REFERENCE 5 (bases 1 to 4947)
AUTHORS Malone,J.A.
JOURNAL Direct Submission
TITLE Submitted (19-AUG-1994) James A. Malone, International Technical Services, Molecular Biology Reagents Division, Pharmacia Biotech Inc., 2202 N. Battelle Ave., Milwaukee, WI 53202-1009, USA
FEATURES
source location/Qualifiers
1..4947
/organism="unidentified cloning vector"
/db_xref="taxon:45196"
/lab_host="Escherichia coli"
183..211
/gene="lac"
183..211
/gene="lac"
/note="lac promoter for inducible expression of glutathione S-transferase"
217..237
/gene="lac"
217..237
/gene="lac"
/note="lac operator"
/bound_moiety="lac repressor protein"
258..959
/citation={3}
/citation={4}
/codon_start=1
/transl_table=11
/product="glutathione S-transferase"
/protein_id="AA57086.1"
/db_xref="GI:595706"
/translation="MSPILGKWKIKGLVQPRLLLEKYEEHLVERDEGDKWRNK
KFELELPEPNLPYIDGDKVLTQSMALIRYIADKHNMLGCGCKREAEISMEGAVLDI
RYGVSRIVAKDFELKVDPLSKPEMLKMFEDRLCHTYLNGDHVTHDPDFLYALD
VLYVMDPMCLAFKPLVCFKKRIEAIPIQIDKYLKSKYIAMPLOGMAMTFGGGDHPK
SDLVKRSPEPIVVD"
918..935
/note="encodes thrombin recognition site"
misc-feature
930..944
/note="Multiple Cloning Site (MCS): contains the unique restriction sites BamHI, SmaI and EcoRI; EcoRI site is in frame with EcoRI site of lambda gtl1"
1285..2215
/gene="bla"
1285..1313
/gene="bla"
1355..2215
/gene="bla"
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="AA57087.1"
/db_xref="GI:595707"
/translation="MSIOHFRVALIPFAAFCLPYFAHPEITVKKDAEDLCARVY
TELDNSKRLIESRPEREFPMSTFFKVLGCAVLSRVADGEOIGRIRHSQNDLVE
YSPVTEKHLTDMIVRELCSAATMSDNTAANLITJIGPEKELAFCHNMDVTRL
DRWEPELNEAIPNDESDTTPAAMAATTKLITGLTILASQOQLIDMWEADKVGPL
LRSDLPAGMFIADKSGAEGSRGIIIALGPGKRSRIIVITYTTSQATMDERNQIA
EIGASLIKHM"
rep_origin
2974
/note="base 2974 represents the first base of the newly synthesized single strand"
gene
3296..4378
/direction=right
/gene="lacIq"
3296..4378
CDS
/gene="lacIq"

BASE COUNT: 1223 a 1193 c 1283 g 1248 t
ORIGIN
Query Match 62.3%; Score 682.4; DB 12; Length 4947;
Best local Similarity 99.1%; Pred. No. 2.1e-165;
Matches 686; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
OY 1 ATGTCCCTATAGTATAGTATGGAATAATGAAGCGCTGTGCAACCACTGCATCTT 60
/codon_start=1
/transl_table=11
/product="lac repressor"
/protein_id="AA57086.1"
/db_xref="GI:595708"
/translation="MKPVTLYDVAYEACVSYQTSRVVYNQASHVSAKPREKYEAMAE
LNTIPRVAQQLAGKOSLLIGVATSSALAHASQIYAATKSRADQIGASVSVSWEBS
GVEACKAAVHNLAAORVSGLLINYPDLDDDAIAVEACTNVPALDLYSDVPINSII
FSHEDGTRGLVEHLVALGHOQIALLGLPSVSAARLRLAGMRHYLRNOIQPIAEREG
DMSASGFOOTQOMLNEGIVPTAMLVANDOMALGMARATIEGRLGADISVGYDDT
ESSCYIPLTTIKODFERLIGOTSVDRILQLSGQAVKGNOLLPLSVLRKRTLLAPNT
OTASPRALADSIMOLAROVSRLESQ"
Db 258 ATGTCCCTATAGTATAGTATGGAATAATGAAGCGCTGTGCAACCACTGCATCTT 317
OY 61 TTGCAATATCTTGAGAAAAATVGAAGACATTTGTATGAGCGCGATGAGGTGATAA 120
Db 318 TTGCAATATCTTGAGAAAAATVGAAGACATTTGTATGAGCGCGATGAGGTGATAA 377
OY 121 TGGCGAACAACAAAGTTTGATGGGTTGGAGTTCCCAATCTCTTATATATGAT 180
Db 378 TGGCGAACAACAAAGTTTGATGGGTTGGAGTTCCCAATCTCTTATATATGAT 437
OY 181 GGTGATGTTAAATTAACACAGCTATGCGCATATGATATAGCTGACAGCAAC 240
Db 438 GGTGATGTTAAATTAACACAGCTATGCGCATATGATATAGCTGACAGCAAC 497
OY 241 ATGTGGGTGGTGTCCAAAGACCGTCGAGAGATTTCAATGCTTGAAGAGCGGTTTG 300
Db 498 ATGTGGGTGGTGTCCAAAGACCGTCGAGAGATTTCAATGCTTGAAGAGCGGTTTG 557
OY 301 GATATATAGTATAGCTTTTGGAGATTTGCATATAGTAAACCTTGAAACTGCCAAAGT 360
Db 558 GATATATAGTATAGCTTTTGGAGATTTGCATATAGTAAACCTTGAAACTGCCAAAGT 617
OY 361 GATTTCTTAGCAAGCTACCTGAATGCTGAATGCTGCAAGTCTTATATGATCAA 420
Db 618 GATTTCTTAGCAAGCTACCTGAATGCTGAATGCTGCAAGTCTTATATGATCAA 677
OY 421 ACATATTTAAATGATGATCAATGATCAATGCTGATTCATGTTATGACGCTTGAT 480
Db 678 ACATATTTAAATGATGATCAATGATCAATGCTGATTCATGTTATGACGCTTGAT 737
OY 481 GTTGTTTTATACATGACCAATATGCTGATGACGCTTCCCAAAATTAAGTTTAA 540
Db 738 GTTGTTTTATACATGACCAATATGCTGATGACGCTTCCCAAAATTAAGTTTAA 797
OY 541 AAACGATTTGAACCTATCCCAAAATGATTAAGTATGAAATCCAGCAAGTATATGCA 600
Db 798 AAACGATTTGAACCTATCCCAAAATGATTAAGTATGAAATCCAGCAAGTATATGCA 857
OY 601 TGGCCTTTGAGAGGCTGCGAAGCCAGCTTGTGTGGCGACCATCTCCAAAATCGAT 660
Db 858 TGGCCTTTGAGAGGCTGCGAAGCCAGCTTGTGTGGCGACCATCTCCAAAATCGAT 917
OY 661 CTGTTCCGCGTGAATCCCGAATTCGCGGAT 692
Db 918 CTGTTCCGCGTGAATCCCGAATTCATGCT 949
RESULT 6
AB014641 3417 bp DNA circular SYN 30-SEP-1998
LOCUS Cloning vector pGEX-PUC-3T DNA, complete sequence.

Services, Molecular Biology Reagents Division, Pharmacia Biotech Inc., 2202 N. Bartlett Ave., Milwaukee, WI 53202-1009, USA

FEATURES

source

1. 4948
Location/Qualifiers

/organism="unidentified cloning vector"

/db_xref="taxon:45196"

/lab_host="Escherichia coli"

183. 211

/standard_name="tac"

/note="tac promoter for inducible expression of

glutathione S-transferase"

217. 237

/gene="lac"

/bound_moiety="lac repressor protein"

217. 237

/gene="lac"

258. 956

/citation={3}

/citation={4}

/codon_start=1

/transl_table=11

/product="glutathione S-transferase"

/protein_id="AA57089.1"

/db_xref="GI:595710"

/translation="MSPILGYWKIKGLVQPTLLLEEKYEHLRYERDEGKRNK

KEFLGEPNLPYIDGDKLKQSMALIRYIADKINMKGGCKEKEAEISMEGLVADI

RYGVSRIRYKDEFTDKLVKFLSKLPEMKLHEDRKLTYLGDHVTLPDPELVYALD

VLYMDPNCIDAFPLVCFKKRIEAIPIQIDTKSKSYIAMPLQGMQATFFGGDHPK

SDLYRGSFGIHRD"

918. 935

/note="encodes thrombin recognition site"

930. 945

/note="Multiple Cloning Site (MCS): contains the unique

restriction sites BamHI, SmaI and EcoRI"

1286. 2216

/gene="bla"

1286. 1314

/gene="bla"

1356. 2216

/gene="bla"

/codon_start=1

/transl_table=11

/product="beta-lactamase"

/protein_id="AA57090.1"

/db_xref="GI:595711"

/translation="MSIOHFRVALIPFAAPCLPYFAHPENTLVKKAEDOLGARVY

IEIDNSKRIEERPERPRPMSTFKVLLCGAVSRVDAGGEOLGRRIHYSQNDLVE

YSPYERKLTDSMTVRELCSAITMSDNTAANLLTTGGRELTAFJHNMGDHVTPL

DRWEPELNEALPNDERTTTPAAMATTLRLKLTGLTLLASRQOLIDWMEADKVGPL

LRSLAPACMFIAIDKSGAGERGRIIAALGPDGKPSRIIVYITGSOATMDERNQIA

EIGASLIKHM"

2974

/note="base 2974 represents the first base of the newly

synthesized single strand"

3297. 4379

/direction="right"

/gene="lacIq"

3297. 4379

/gene="lacIq"

/codon_start=1

/transl_table=11

/product="lac repressor"

/protein_id="AA57091.1"

/db_xref="GI:595712"

Query Match 62.1%; Score 681; DB 12; Length 4948;

Best Local Similarity 100.0%; Pred. No. 4,9e-165; Mismatches 0; Indels 0; Gaps 0;

Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGTCCTATACCTAGGATTTGGAAATTTAGAGGCTTGTGACACCACTGCTTCTT 60

258 ATGTCCTATACCTAGGATTTGGAAATTTAGAGGCTTGTGACACCACTGCTTCTT 317

61 TTGGAATATCTTGAAGAAAAATATGAAGACATTTGTATAGCGCATGAAGTGTAA 120

318 TTGGAATATCTTGAAGAAAAATATGAAGACATTTGTATAGCGCATGAAGTGTAA 377

121 TGGCGAACAACAAAGCTTTGCAATGGCTTGGAGTTTCCCAATCTTCTTATATATGAT 180

378 TGGCGAACAACAAAGCTTTGCAATGGCTTGGAGTTTCCCAATCTTCTTATATATGAT 437

181 GGTGATGTTAAATTAACACAGTCTATGACATCATGTTATATACCTGACACACAC 240

438 GGTGATGTTAAATTAACACAGTCTATGACATCATGTTATATACCTGACACACAC 497

241 ATGTTGGGTGGTGTGTCACAAAGAGCTGACAGATTTCAATGCTTGAAGAGCGGTTTG 300

498 ATGTTGGGTGGTGTGTCACAAAGAGCTGACAGATTTCAATGCTTGAAGAGCGGTTTG 557

301 GATATTAGATACGCTGTTGAGAAATTTGCAATGATTAAGACCTTGAACCTGCAAGTT 360

558 GATATTAGATACGCTGTTGAGAAATTTGCAATGATTAAGACCTTGAACCTGCAAGTT 617

361 GATTTTCTTAGCAAGCTACCTGAATGCTGAATTTGCAATGCTTGAATGCTTGAAT 420

618 GATTTTCTTAGCAAGCTACCTGAATGCTGAATTTGCAATGCTTGAATGCTTGAAT 677

421 ACATATTAAATGATGATCATGTAACCATCTGACTTCAATGTTATGACGCTTTGAT 480

678 ACATATTAAATGATGATCATGTAACCATCTGACTTCAATGTTATGACGCTTTGAT 737

481 GTTGTTTATACATGACCAATGTCCTGATGCGTTCGCCAATTAAGTTGTTTAA 540

738 GTTGTTTATACATGACCAATGTCCTGATGCGTTCGCCAATTAAGTTGTTTAA 797

541 AAACGATTTGAACCTATCCCAAAATTTGATTAAGTCTTGAATCCAGCAATATATAGCA 600

798 AAACGATTTGAACCTATCCCAAAATTTGATTAAGTCTTGAATCCAGCAATATATAGCA 857

601 TGGCCTTTGACAGGCTGCGCAAGCCAGCTTTGTTGTTGCGGACCATCTCCAAAATCGGAT 660

858 TGGCCTTTGACAGGCTGCGCAAGCCAGCTTTGTTGTTGCGGACCATCTCCAAAATCGGAT 917

661 CTGCTCCGCGGTGATCCCG 681

918 CTGCTCCGCGGTGATCCCG 938

RESULT 8

EV084572 5042 bp DNA circular SYN 04-FEB-1997

LOCUS EV084572 Expression vector pGEX-2T-6His-PL2 (GST) gene, complete

DEFINITION cds, complete vector sequence.

ACCESSION 084572

VERSION 084572.1 GI:1814367

KEYWORDS

SOURCE

ORGANISM

Expression vector pGEX-2T-6His-PL2.

Expression vector pGEX-2T-6His-PL2.

Artificial sequence; vectors.

REFERENCE

1 (bases 1 to 5042)

Hipskind, R.H. and Delaney, P.

Unpublished

JOURNAL

2 (bases 1 to 5042)

REFERENCE

Cahill, M.A.

Direct Submission

TITLE

Direct Submission

Direct Submission

Direct Submission

Direct Submission

Direct Submission

Direct Submission

Direct Submission

Direct Submission

Direct Submission

Direct Submission

Direct Submission

Direct Submission

BASE COUNT

1223 a 1193 c 1284 g 1248 t

ORIGIN

Db 438 GGTGATGTTAAATTAACACAGCTATGCGCATCATCGTTATATAGCTGACACACAC 497
 Qy 241 ATGTGGGATGTTGTCACAAAGAGCGCTGACAGATTTCAATGCTTGAAGAGCGGTTTGG 300
 Db 498 ATGTTGGGATGTTGTCACAAAGAGCGCTGACAGATTTCAATGCTTGAAGAGCGGTTTGG 557
 Qy 301 GATATTAGTACGCGTGTTCGAGAAATGTCATATAGTAAGACTTTGAAACTCTCAAAATT 360
 Db 558 GATATTAGTACGCGTGTTCGAGAAATGTCATATAGTAAGACTTTGAAACTCTCAAAATT 617
 Qy 361 GATTTTCTAGACAGCTACCTGAAATGTCGAAATGTCGAAAGATGCTTATGCTCAATAA 420
 Db 618 GATTTTCTAGACAGCTACCTGAAATGTCGAAATGTCGAAAGATGCTTATGCTCAATAA 677
 Qy 421 ACATATTTAAATGTCGATCATGTAAACCATCTGACTTCATGTTATGACCGCTTTGAT 480
 Db 678 ACATATTTAAATGTCGATCATGTAAACCATCTGACTTCATGTTATGACCGCTTTGAT 737
 Qy 481 GTTGTGTTTATACATGACCAATGTCCTGATGCGTTCGCCAAATATAGTTGTTTAA 540
 Db 738 GTTGTGTTTATACATGACCAATGTCCTGATGCGTTCGCCAAATATAGTTGTTTAA 797
 Qy 541 AAACGATTTAGAGCTATGTCACAAATGATATAGTACTTGAATCCGACAGATATATAGA 600
 Db 798 AAACGATTTAGAGCTATGTCACAAATGATATAGTACTTGAATCCGACAGATATATAGA 857
 Qy 601 TGGCCCTTTCAGAGCGCTGACAGCGATGTTGCTGCTGCGGACCATCTCCAAATATGAT 660
 Db 858 TGGCCCTTTCAGAGCGCTGACAGCGATGTTGCTGCTGCGGACCATCTCCAAATATGAT 917
 Qy 661 CTGCTTCGCGGTGATCCCG 681
 Db 918 CTGCTTCGCGGTGATCCCG 938

RESULT 10

LOCUS EV085201 5501 bp DNA circular SYN 31-MAR-1998
 DEFINITION Expression vector pGH/F2.2-163, complete sequence.
 ACCESSION U85201
 VERSION U85201.1 GI:3002495

KEYWORDS Expression vector pGH/F2.2-163.
 ORGANISM Expression vector pGH/F2.2-163.
 SOURCE Expression vector pGH/F2.2-163.
 REFERENCE 1 (bases 1 to 5501)
 AUTHORS Cahill, M.A.
 TITLE Phosphorylation of Transcription Factors Elk-1 and Fra-2 by Associated Serum-dependent Kinase Activity

JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 5501)
 AUTHORS Cahill, M.A.
 TITLE Direct Submission

JOURNAL Submitted (14-JAN-1997) Division of Immunology & Cell Biology,
 JCSMR, Australia National University, P.O. Box 334, Canberra, ACT
 2601, Australia

FEATURES
 source location/Qualifiers
 1..5501

gene /organism="Expression vector pGH/F2.2-163"
 /db_xref="taxon:56100"
 /note="pGEX-2t derivative"
 258..1475
 /gene="GST/F2.2-163"
 258..1475
 /gene="GST/F2.2-163"
 /codon_start=1
 /transl_table=11
 /product="GSTmF2a/2-163"
 /protein_id="AAC08715.1"
 /db_xref="GI:3002495"
 CDS
 /translation="MSPILGYWKIKGIQVOPFLLEYLEEKYEELHEDDEKRRNK
 KRELLEPPNLPYITDGVRKLTQSAITRIYIDKINMLGGCKEKAELISMEGAVLDI
 RGVSRIVASIDFETLKVDFLSKLPKMLKEDRLCHRTYINGDHVTHDEPMLDALD

VVLVDPKCLDAEPKLVCFKRIEAIPOIDKYLKSKXYIAMPLOGMATFGGDDHPK
 SDLVPRGSPVSHHHHHHGEKILYODPGNFDSSRSGSSPAHSEYSSGGGQQ
 KRVPMPGSAFIPTIAITTSODLQMWQPTVITSMNPPRSHVPSLPLGLASVP
 GHALPRLPVITIGTITGVRRRRDQLSPHEEEKRIRRENRKLAACORNRRELTE
 KIOAETEELEEK"
 1909..2769
 /codon_start=1
 /transl_table=11
 /product="beta-lactamase"
 /protein_id="AAC08716.1"
 /db_xref="GI:3002497"
 /translation="MSIQHFRVALIPEFAFLPYPAHEPILVKKVDAEDQIARVY
 IELDINSKILFESFRPEERFPMSTFEKVLICGAVLSRDVAGQDGLRIRHYSQNDLVE
 YSPVEKHLTDMTVRELCSAAITMSDNTAAAILLTITIGPPELTAFLIMAGDHYRL
 DWPEPELNEAIIPNDERDQITMPPAAMATLITGELTGLISRAQOIIDMEKAVKVGPL
 LRSALPAGWFIADKSGAGERSRGIIAALPGPKRSRIIVITYTSSQATMDERNQIA
 ELGASLIKHW"
 3841..4932
 /codon_start=1
 /transl_table=11
 /product="lac operon repressor"
 /protein_id="AAC08717.1"
 /db_xref="GI:3002498"
 /translation="WVNVKPVLYVAEYAGVSYOTVSHVNVQASHVSKTREKVEAA
 NAEIVIPNRVQQLAKQSLIGVATSLAHSQIVAAIKSRADQIGASVSMV
 ERSQVEACKAAVHNLDAORVSLIINYPLDDQDAIYAEEACNIVPLAFIDQPIIN
 SIFSHEDQTRGLVHVALGHQOIALAGPLSSVARSRLRGMKHYLFRNODIPAE
 RSGDSAMSQFOOTOMINEGIPTVPMVAIYNOMALGNARATTEGSLRPGADISVYG
 DTEBDSCTIPEPLTTIKODFRILGOTVDRILQLSOGQAVKRNQILPVLVRRKTTLLA
 PMTQTPASPALDSLMQARVSRSESG"

CDS

BASE COUNT 1353 a 1367 c 1444 g 1337 t
 ORIGIN

Query Match 62.1%; Score 681; DB 12; Length 5501;
 Best Local Similarity 100.0%; Pred. No. 5e-165;

Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTCCCTATACTAGTATTTGAAATTAAGGCGCTTGCACCCACTGCATCTTCT 60
 Db 258 ATGTCCCTATACTAGTATTTGAAATTAAGGCGCTTGCACCCACTGCATCTTCT 317
 Qy 61 TTGGAATATCTTGAAGAAAATATGAAGACATTTGATAGCGGCGATGAAGTGATAA 120
 Db 318 TTGGAATATCTTGAAGAAAATATGAAGACATTTGATAGCGGCGATGAAGTGATAA 377
 Qy 121 TGGCGAACAACAAAGTTGATTTGGGTTGGAGTTTCCCATCTTCTTATATTTGAT 180
 Db 378 TGGCGAACAACAAAGTTGATTTGGGTTGGAGTTTCCCATCTTCTTATATTTGAT 437
 Qy 181 GGTGATGTTAAATTAACACAGCTATAGGCCATCATCGTATATAGCTGACAGCAAC 240
 Db 438 GGTGATGTTAAATTAACACAGCTATAGGCCATCATCGTATATAGCTGACAGCAAC 497
 Qy 241 ATGTGGGATGTTGTCACAAAGAGCGCTGACAGATTTCAATGCTTGAAGAGCGGTTTGG 300
 Db 498 ATGTGGGATGTTGTCACAAAGAGCGCTGACAGATTTCAATGCTTGAAGAGCGGTTTGG 557
 Qy 301 GATATTAGTACGCGTGTTCGAGAAATGTCATATAGTAAGACTTTGAAACTCTCAAAATT 360
 Db 558 GATATTAGTACGCGTGTTCGAGAAATGTCATATAGTAAGACTTTGAAACTCTCAAAATT 617
 Qy 361 GATTTTCTAGACAGCTACCTGAAATGTCGAAATGTCGAAAGATGCTTATGCTCAATAA 420
 Db 618 GATTTTCTAGACAGCTACCTGAAATGTCGAAATGTCGAAAGATGCTTATGCTCAATAA 677
 Qy 421 ACATATTTAAATGTCGATCATGTAAACCATCTGACTTCATGTTATGACCGCTTTGAT 480
 Db 678 ACATATTTAAATGTCGATCATGTAAACCATCTGACTTCATGTTATGACCGCTTTGAT 737
 Qy 481 GTTGTGTTTATACATGACCAATGTCCTGATGCGTTCGCCAAATATAGTTGTTTAA 540
 Db 738 GTTGTGTTTATACATGACCAATGTCCTGATGCGTTCGCCAAATATAGTTGTTTAA 797

SOURCE Expression vector pGH/F2.159-327.
ORGANISM Expression vector pGH/F2.159-327
REFERENCE 1 (bases 1 to 5516)
AUTHORS Cahill, M.A.
TITLE Phosphorylation of Transcription Factors Elk-1 and Fra-2 by Associated Serum-dependent Kinase Activity
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 5516)
AUTHORS Cahill, M.A.
TITLE Direct Submission
JOURNAL Submitted (14-JAN-1997) Division of Immunology & Cell Biology, JCSMR, Australia National University, P.O. Box 334, Canberra, ACT 2601, Australia

FEATURES
source 1. 5516
/organism="Expression vector pGH/F2.159-327"
/db_xref="taxon:36102"
/note="pGEX-2T derivative"
258. 1490
/gene="GST/F2.159-327"
258. 1490
/gene="GST/F2.159-327"
/codon_start=1
/transl_table=1
/product="GstF2.159-327"
/protein_id="AAC08721.1"
/db_xref="GI:3002504"
/translation="MSPTLGGKIKGLVOPTRLLLEYLEEKYEHLERDEGDKRNRK
KFEJLEPNNLPYIDGVKLTQSMALLRYIADKRNMGCCKEAEISMEGAVLDI
RYGVSRVAYSDPEFTLVKDFLSKLEPMKMEEDRLCHKTYLNGDHVTHPEMLADALD
VLYWMDPMLDAFPLKVCFKRKEIAIPQIDKYLKSKYIAMLQGMQATFGGDHPK
SDLVGRSPVMSHHHHEFEKLEEKSGLOKRIAELOKEKELEEMKAVAGPVCK
ISPERRSPPSGLOSLRGTSAGVPPVVKQPEDEPSSAGMDKORVYKPKISI
AGGFGTEPPLHTPIVYVSTPAITPGTINLVFTTNYVLEQESPSSPSECSKAHRRSS
SSGDSDSLNSPTLLAL"
1924. 2784
/codon_start=1
/transl_table=1
/product="beta-lactamase"
/protein_id="AAC08723.1"
/db_xref="GI:3002506"
/translation="MSIOHFRVALIPFEAFCLPVFAHPELVKVKDAEDQLGARVY
IELDLSNGKILSEPERFPMSTFKVLGCVGSRVADGEOLGRRIHVSQNDLVE
YSPTEKHLTDGMTVELCSAITMSDNTANLLITIGSPETLAPFLHMGDIYTRL
DMREPELNEATPNDREDTMPRAAMTTRKLLTGLLILASNOQLIDMWEADKVGPL
LRSLPAGMTIADKSGAGERSGIIIALGPDGKPSRIIVITTSQATMDERNOIA
EIGASLIRHW"
3856. 4947
/codon_start=1
/transl_table=1
/product="lac operon repressor"
/protein_id="AAC08722.1"
/db_xref="GI:3002505"
/translation="MVNKKPYTLVDVAEYAGYQTVSRVNVQASHVSAKPREKVEAA
MAELNYPINRAOOLAKOSLILGATSSLSLHAISOYIAALKSRADOLGASVVSMV
ERSGVEACKAAPHNLAAORVSGLIINYPDDODATAVAACACNVAPALIDVSDOPIN
SIIRSHEDGTRLGEVHLYALGHQDIALLAGSIVSSARLRAGHKRYLTRNDQIPAE
REGDWSASGFOQTQMOLNEGIVPTAMLVANDOMLVGMRATTEGLRVGADISVVG
DDEDSGCIYPLPTIKODPFLIGOTSVDRLOLSQCAVKNOLLPSLVKRRKTTLLA
PNTQTPASPRALADSLMOLAROVSRLESGO"

BASE COUNT 1348 a 1379 c 1440 g 1349 t
ORIGIN

Query Match 62.1%; Score 681; DB 12; Length 5516;
Best Local Similarity 100.0%; Pred. No. 5e-165;
Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCCCTATACCTAGTATTGGAATAATTAGGCGCTTGACACCCACTGCATCTCT 60
|||||
Db 258 ATGTCCCTATACCTAGTATTGGAATAATTAGGCGCTTGACACCCACTGCATCTCT 317

QY 61 TTGGAATATCTTGAGAAAAATATGAGAGCATTTGTATGAGCGCGATGAGTGATAAA 120
|||||
Db 318 TTGGAATATCTTGAGAAAAATATGAGAGCATTTGTATGAGCGCGATGAGTGATAAA 377
|||||
QY 121 TGGCGAACAACAAAGCTTGAATTGGGTTGGAGTTTCCCAATCTTCTTATATATGAT 180
|||||
Db 378 TGGCGAACAACAAAGCTTGAATTGGGTTGGAGTTTCCCAATCTTCTTATATATGAT 437
|||||
QY 181 GGTGATGTTAAATTAACACAGTCTATGCGCATCATCGTTATATAGCTGACAGCAAC 240
|||||
Db 438 GGTGATGTTAAATTAACACAGTCTATGCGCATCATCGTTATATAGCTGACAGCAAC 497
|||||
QY 241 ATGTGGGTTGGTGTTCACAAAGAGCGTGCAGAGATTTCATATGCTTGAAGAGCGGTTTG 300
|||||
Db 498 ATGTGGGTTGGTGTTCACAAAGAGCGTGCAGAGATTTCATATGCTTGAAGAGCGGTTTG 557
|||||
QY 301 GATATTAGTACGCGTTTTCGAGAAATTGCATATAGTAAGACTTTAAACTCCTCAAGTT 360
|||||
Db 558 GATATTAGTACGCGTTTTCGAGAAATTGCATATAGTAAGACTTTAAACTCCTCAAGTT 617
|||||
QY 361 GATTTCTAGCAAGCTACCTGAAATGCTGAAATGCTGAAAGATGCTTATGCTATAA 420
|||||
Db 618 GATTTCTAGCAAGCTACCTGAAATGCTGAAATGCTGAAAGATGCTTATGCTATAA 677
|||||
QY 421 ACATATTTAAATGCTGATCATATGTAACCATCTGACTGATGTTATGACGCTTTGAT 480
|||||
Db 678 ACATATTTAAATGCTGATCATATGTAACCATCTGACTGATGTTATGACGCTTTGAT 737
|||||
QY 481 GTTGTTTTATCATGACCAATGTCCTGATGCTGCTGCCAAATATGTTGTTTAAA 540
|||||
Db 738 GTTGTTTTATCATGACCAATGTCCTGATGCTGCTGCCAAATATGTTGTTTAAA 797
|||||
QY 541 AAACGATTGAAGCTATCCACAAATTTGATTAAGTAAATCCGCAAGTATATAGCA 600
|||||
Db 798 AAACGATTGAAGCTATCCACAAATTTGATTAAGTAAATCCGCAAGTATATAGCA 857
|||||
QY 601 TGGCCTTTGACAGGCTGCGCAAGCCAGCTTGTGTGTGCGCACCATCTCCAAATGCGAT 660
|||||
Db 858 TGGCCTTTGACAGGCTGCGCAAGCCAGCTTGTGTGTGCGCACCATCTCCAAATGCGAT 917
|||||
QY 661 CTGGTTCGCGGTGATCCCG 681
|||||
Db 918 CTGGTTCGCGGTGATCCCG 938
|||||

RESULT 13
EV085204
LOCUS EV085204 5738 bp DNA circular SYN 31-MAR-1998
DEFINITION Expression vector pGH/F2.2-242, complete sequence.
ACCESSION U85204
VERSION U85204.1 GI:3002507
KEYWORDS
SOURCE
ORGANISM Expression vector pGH/F2.2-242.
REFERENCE 1 (bases 1 to 5738)
TITLE Phosphorylation of Transcription Factors Elk-1 and Fra-2 by Associated Serum-dependent Kinase Activity
AUTHORS Cahill, M.A.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 5738)
TITLE Direct Submission
AUTHORS Cahill, M.A.
JOURNAL Submitted (14-JAN-1997) Division of Immunology & Cell Biology, JCSMR, Australia National University, P.O. Box 334, Canberra, ACT 2601, Australia

FEATURES
source 1. 5738
/organism="Expression vector pGH/F2.2-242"
/db_xref="taxon:36103"
/note="pGEX-2T derivative"
258. 1712
/gene="GST/F2.2-242"

/protein_id="AAC08729.1"
 /db_xref="GI:3002514"
 /translation="MSIOHFRVALLIPFEARCLPYFAHPETLVKKAEDOLGARVGY
 IELDLSNGKILSEFRPEERPMSTFKVLGCAVLSRVNADGEOLGRTHYSONDLVE
 YSPVTEKHLTDGMTVRELCSSAITSNDNTAANLLITPGPELTAFLNNKGDHVTRE
 DRWPEELNEATIPNDRDTPPMAMATTLRLKLTGELTLASRQOLIDMKADKVAQPL
 LRSALPAGMFIADKSGAGERSRGITIAALGPDGPKSRIVIVYTTQSOATMDERNRQIA
 EIGASLIKHW"
 4096..5187
 /codon_start=1
 /product="lac operon repressor"
 /protein_id="AAC08728.1"
 /db_xref="GI:3002513"
 /translation="MVNVKPYTLVDVAEYAGVSYQTVSRVNVQASVSAKREKVEAA
 MAELNYIPNRVAQOLAGKOSLLIGVATSSLAHARQIYAALIKSRDOLGASVYVSM
 ERSQVACKAAVHNLAORVSGLLINYPDDDAVAEACNTNPALFLDVSDDPFIN
 SLTSHEDGTRLGVEHLVALGHOQIALLAGPLSSASARLRAGMKRLTRNDIOTPAE
 REGDMSASGROOTMOMNEGIVPTAMLVANDOMALGAMRLTESGLRVGADISVGY
 DQTESSCYIPPLTTIKODPFRLLGQTSVDRLLQLISOGAVKGNOLLPSVLVKKRTTLA
 PNTQTSAPRALDSIMOLAROVSRLESGO"

BASE COUNT 1407 a 1448 c 1520 g 1381 t
 ORIGIN

Query Match 62.1%; Score 681; DB 12; Length 5756;
 Best Local Similarity 100.0%; Pred. No. 5e-165;
 Matches 681: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCCCTTAACCTAGCTTATTGGAATAATTAGGCGCTTGTCACACCCACTGCATCTT 60
 DB 258 ATGTCCCTTAACCTAGCTTATTGGAATAATTAGGCGCTTGTCACACCCACTGCATCTT 317
 QY 61 TTGAATATCTGAAGAAAATATGAGACATTTGTATGAGCGGCGATGAGGTGATAA 120
 DB 318 TTGAATATCTGAAGAAAATATGAGACATTTGTATGAGCGGCGATGAGGTGATAA 377
 QY 121 TGGCGAACAACAAAGTTGGAATTTGGTTGAGTTTCCCAATCTTCCTATTATTATGAT 180
 DB 378 TGGCGAACAACAAAGTTGGAATTTGGTTGAGTTTCCCAATCTTCCTATTATTATGAT 437
 QY 181 GGTGATGTTAAATTAACACAGCTCTATAGCCATCATCGTTATATAGCTGACACACAC 240
 DB 438 GGTGATGTTAAATTAACACAGCTCTATAGCCATCATCGTTATATAGCTGACACACAC 497
 QY 241 ATGTGGGTGGTTGCCAAGAAGCGTGCAGAGATTCAATGCTTGAGAGAGCGGTTTG 300
 DB 498 ATGTGGGTGGTTGCCAAGAAGCGTGCAGAGATTCAATGCTTGAGAGAGCGGTTTG 557
 QY 301 GATATTAGATACGCTGTTTCGAGAAATTGCATATAGTAAGACTTTGAAACTCTCAAGTT 360
 DB 558 GATATTAGATACGCTGTTTCGAGAAATTGCATATAGTAAGACTTTGAAACTCTCAAGTT 617
 QY 361 GATTTTCTTACAGAGCTACCTGAATGCTGAANAATTTGAGAGATGCTTATATGCTATAA 420
 DB 618 GATTTTCTTACAGAGCTACCTGAATGCTGAANAATTTGAGAGATGCTTATATGCTATAA 677
 QY 421 ACATATTTAAATGSGTATGATGTAACCACTCCGACTTCATGTTGATGAGCGCTTGAT 480
 DB 678 ACATATTTAAATGSGTATGATGTAACCACTCCGACTTCATGTTGATGAGCGCTTGAT 737
 QY 481 GTTGTGTTTATACATGAGCAACATGTCCTGATGCGTCCCAAAATTAAGTTGTTTAA 540
 DB 738 GTTGTGTTTATACATGAGCAACATGTCCTGATGCGTCCCAAAATTAAGTTGTTTAA 797
 QY 541 AAACGATTGAGAGCTATCCGACAATTTGATAGTACTGTAATCCGACAGATATATAGCA 600
 DB 798 AAACGATTGAGAGCTATCCGACAATTTGATAGTACTGTAATCCGACAGATATATAGCA 857
 QY 601 TGGCCTTTTGCAGAGGCTGGCAGACGCTTGTGTTGGGAGCATCCTCCAAAATGCGAT 660
 DB 858 TGGCCTTTTGCAGAGGCTGGCAGACGCTTGTGTTGGGAGCATCCTCCAAAATGCGAT 917

QY 661 CTGCTTCCCGCTGATCCCG 681
 DB 918 CTGCTTCCCGCTGATCCCG 938

RESULT 15

EV085206 5990 bp DNA circular SYN 31-MAR-1998
 DEFINITION Expression vector pGH/F2.2-327, complete sequence.
 ACCESSION U85206
 VERSION U85206.1 GI:3002515
 KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

REGWSAMSGFOQTMOMINEGIVPTAMLVANDOMALGAMRAITTESGLRVGADISVVG
 DDEEDSSCYIPILPTIKODFRILGQSVDRILQLSOGQAVKGNQLPLVSLVKRTTLLA
 PNCQTASPPALDSIMOLAROVSRLESQ"
 BASE COUNT 1454 a 1532 c 1580 g 1424 t
 ORIGIN

Query Match 62.1%; Score 681; DB 12; Length 5990;
 Best Local Similarity 100.0%; Pred. No. 5e-165;
 Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCCCTATACAGGATTTTGAAGATTTAGGGCTTGTGCAACCCACTGACTTCTT 60
 |||||||
 Db 258 ATGTCCCTATACAGGATTTTGAAGATTTAGGGCTTGTGCAACCCACTGACTTCTT 317
 QY 61 TTGGAATATCTTGAAGAAAAATATGAGAGCATTTGTATAGCGCGATGAGGTATAA 120
 |||||||
 Db 318 TTGGAATATCTTGAAGAAAAATATGAGAGCATTTGTATAGCGCGATGAGGTATAA 377
 QY 121 TGGCGAACAACAAAGTTTGAATGGGTTTGGAGTTTCCCAATCTTCTTATTATTTGAT 180
 |||||||
 Db 378 TGGCGAACAACAAAGTTTGAATGGGTTTGGAGTTTCCCAATCTTCTTATTATTTGAT 437
 QY 181 GGTGATGTAAATTAACACAGCTATGCGCATCATACGTTATATAGCTGACAGACAAC 240
 |||||||
 Db 438 GGTGATGTAAATTAACACAGCTATGCGCATCATACGTTATATAGCTGACAGACAAC 497
 QY 241 ATGTGGGTGTTGTCACAAAGAGCGTGACAGATTTCAATGCTTGAAGAGCGGTTTG 300
 |||||||
 Db 498 ATGTGGGTGTTGTCACAAAGAGCGTGACAGATTTCAATGCTTGAAGAGCGGTTTG 557
 QY 301 GATATTAGATACGGTGTGAGAAATGCAATATAGTAAGACTTTGAACCTCTCAAGTT 360
 |||||||
 Db 558 GATATTAGATACGGTGTGAGAAATGCAATATAGTAAGACTTTGAACCTCTCAAGTT 617
 QY 361 GATTTTCTAGCAGCTACCTGAATGCGAAATGTCGAGATCGTTTATGTCATAAA 420
 |||||||
 Db 618 GATTTTCTAGCAGCTACCTGAATGCGAAATGTCGAGATCGTTTATGTCATAAA 677
 QY 421 ACATATTTAAATGTCATCATGTAACCCATCCTGACTTCATGTGTATGACGCTTGAT 480
 |||||||
 Db 678 ACATATTTAAATGTCATCATGTAACCCATCCTGACTTCATGTGTATGACGCTTGAT 737
 QY 481 GTTGTTTATACATGACCCCAATGTCCTGATGCGTTCGCCAAATAGTTTGTAA 540
 |||||||
 Db 738 GTTGTTTATACATGACCCCAATGTCCTGATGCGTTCGCCAAATAGTTTGTAA 797
 QY 541 AAACGATTTGAAGCTATCCCAATGTATAGTACTTGAATCCAGCAAGTATATAGCA 600
 |||||||
 Db 798 AAACGATTTGAAGCTATCCCAATGTATAGTACTTGAATCCAGCAAGTATATAGCA 857
 QY 601 TGGCCTTTGCAAGGGCTGGCAAGCCACGTTTGGTGGGACCATCTCCAAATCGGAT 660
 |||||||
 Db 858 TGGCCTTTGCAAGGGCTGGCAAGCCACGTTTGGTGGGACCATCTCCAAATCGGAT 917
 QY 661 CTGCTTCCGCGTGGATCCCG 681
 |||||||
 Db 918 CTGCTTCCGCGTGGATCCCG 938

Search completed: March 18, 2001, 04:57:19
 Job time: 4243 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2001, 04:17:41 ; Search time 83.44 Seconds
(without alignments)
4934.400 Million cell updates/sec

Title: US-09-402-488a-1

Perfect score: 1096
Sequence: 1 ATGTCCCTACTAGAGTTA.....ATCCACTAGTTAGACCGG 1096

Scoring table: IDENTITY NUC
Gapop 10-0, Gapext 1.0

Searched: 480022 seqs, 187831343 residues

Total number of hits satisfying chosen parameters: 960044

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
N.Geneseq_36: *
1: /cgn2_2/gcgdata/geneseq/geneseqn/NA1980.DAT: *
2: /cgn2_2/gcgdata/geneseq/geneseqn/NA1981.DAT: *
3: /cgn2_2/gcgdata/geneseq/geneseqn/NA1982.DAT: *
4: /cgn2_2/gcgdata/geneseq/geneseqn/NA1983.DAT: *
5: /cgn2_2/gcgdata/geneseq/geneseqn/NA1984.DAT: *
6: /cgn2_2/gcgdata/geneseq/geneseqn/NA1985.DAT: *
7: /cgn2_2/gcgdata/geneseq/geneseqn/NA1986.DAT: *
8: /cgn2_2/gcgdata/geneseq/geneseqn/NA1987.DAT: *
9: /cgn2_2/gcgdata/geneseq/geneseqn/NA1988.DAT: *
10: /cgn2_2/gcgdata/geneseq/geneseqn/NA1989.DAT: *
11: /cgn2_2/gcgdata/geneseq/geneseqn/NA1990.DAT: *
12: /cgn2_2/gcgdata/geneseq/geneseqn/NA1991.DAT: *
13: /cgn2_2/gcgdata/geneseq/geneseqn/NA1992.DAT: *
14: /cgn2_2/gcgdata/geneseq/geneseqn/NA1993.DAT: *
15: /cgn2_2/gcgdata/geneseq/geneseqn/NA1994.DAT: *
16: /cgn2_2/gcgdata/geneseq/geneseqn/NA1995.DAT: *
17: /cgn2_2/gcgdata/geneseq/geneseqn/NA1996.DAT: *
18: /cgn2_2/gcgdata/geneseq/geneseqn/NA1997.DAT: *
19: /cgn2_2/gcgdata/geneseq/geneseqn/NA1998.DAT: *
20: /cgn2_2/gcgdata/geneseq/geneseqn/NA1999.DAT: *
21: /cgn2_2/gcgdata/geneseq/geneseqn/NA2000.DAT: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1096	100.0	1096	20	V83966 DNA encoding the f
2	685	62.5	720	20	X01312 GST coding sequenc
3	684.2	62.4	3591	21	Z58406 DNA encoding funon
4	684.2	62.4	3591	21	Z60645 DNA encoding a fus
5	684	62.4	2475	21	Z32916 Human HCP/GST fusi
6	683.4	62.3	1866	21	Z32917 Human HCP phosphod
7	683.2	62.3	2079	21	Z58400 DNA encoding GST-t
8	683.2	62.3	2079	21	Z60639 DNA encoding an am
9	683	62.3	2490	21	Z58407 DNA encoding GST-t
10	683	62.3	2490	21	Z60646 DNA encoding a fus
11	682.2	62.2	3618	21	Z58405 DNA encoding funon
12	682.2	62.2	3618	21	Z60644 DNA encoding a fus

13	681	62.1	699	16	Q86774 Glutathione-s-tran
14	681	62.1	1695	16	T05329 DNA encoding TCR27
15	681	62.1	1932	16	T05332 DNA encoding TCR27
16	681	62.1	3175	20	Z20066 DNA encoding gluta
17	681	62.1	3867	20	Z20067 DNA encoding gluta
18	679.4	62.0	693	16	O87508 Glutathione-s-tran
19	678.4	61.9	1281	16	T03457 Turkey prolactin/G
20	678.2	61.9	1740	21	Z58065 CAP-A-B fusion pro
21	678	61.9	816	13	Q22843 CAP-A-B fusion pro
22	677.8	61.8	1065	16	T05330 DNA encoding TCR27
23	677	61.8	1119	18	T90496 GST-Treponema pall
24	675.2	61.6	1335	13	Q22855 pGEX-2T-15:17 enco
25	674	61.5	759	13	Q22840 pGEX-2T-CAP-A enco
26	674	61.5	3423	16	T00771 GST-SBP Fusion gen
27	673.6	61.5	1194	16	O86776 GST-SBP-1 gene fus
28	673.2	61.4	759	13	Q22841 pGEX-2T-CAP-B enco
29	673.2	61.4	759	13	Q22842 pGEX-2T-CAP-C enco
30	667.4	60.9	897	16	O82895 Sequence encoding
31	663.4	60.5	751	19	V61099 Plasmid pGEX-PH 75
32	663.4	60.5	1128	16	Q95552 GST-calmodulin fus
33	663.4	60.5	1239	18	T93979 DNA for oligopepti
34	663.4	60.5	1419	16	T05333 DNA encoding TCR27
35	663.4	60.5	1830	19	V61100 Plasmid pGEX-PH 18
36	663.4	60.5	1820	20	X77118 GS-GC6 fusion prot
37	662.8	60.5	2095	20	V55614 pGEX-MSP I methylas
38	661.8	60.4	951	17	T30371 pGEX/hp57 plasmid
39	661.8	60.4	1842	17	T30370 GST/truncated huma
40	661.8	60.4	2067	17	T30369 GST/human p57 fusi
41	661	60.3	924	16	T05331 DNA encoding TCR27
42	660.4	60.3	1851	19	V13177 GST-cyclin E-PEST*
43	659.6	60.2	894	13	Q22852 pGEX-3X-693:691 en
44	658.6	60.1	858	13	Q22853 pGEX-3X-15:18 enco
45	658.6	60.1	948	13	Q22839 pGEX-3X-690:694 en

ALIGNMENTS

RESULT 1	
ID V83966	V83966 standard; DNA; 1096 BP.
XX AC V83966;	
XX DT 09-MAR-1999 (first entry)	
XX DE DNA encoding the fusion protein GST-chymosin pro-peptide-Hirudin.	
XX XX	
KW Fusion protein; bovine chymosin pro-peptide; leech; hirudin;	
KW anticoagulant protein; autocatalytically maturing zymogen;	
KW carp growth hormone; vaccine; ss.	
XX XX	
OS Synthetic.	
OS Bos sp.	
OS Hirudo medicinalis.	
XX XX	
FH Key	Location/Qualifiers
FT CDS	1..1032
FT	/tag= a
FT	/product= GST-chymosin pro-peptide-hirudin
FT	1..708
FT	/tag= b
FT	/note= "encodes Gsn"
FT	709..834
FT	/tag= c
FT	/note= "encodes bovine chymosin pro-peptide"
FT	835..1029
FT	/tag= d
FT	/note= "encodes leech hirudin; sequence has been
XX XX	optimised for plant codon usage"
XX	W09849326-A1.

Query Match	100.0%	Score 1096;	DB 20;	Length 1096;
Best Local Similarity	100.0%;	Pred. No. 3.8e-292;		
Matches 1096;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
0Y 1 ATGTCCCTATCTAGTATTGGAATAATTAAAGGCGCTTGTCACCCACTGCACCTCTT 60				
Db 1 atgtccctatactagtgtatctggaataatlaaggcgcttgcacccactgcactctt 60				
0Y 61 TTGGATATCTTGAAGAAAATATGACAGCATTTGTATGACCGGATGAAGTATATA 120				
Db 61 ttggatatcttgaagaaaatatgaaagacatttgtatgacgcggatgaagtataata 120				
0Y 121 TGGCGAACAACAAAGCTTGAATTTGGGTTTGAAGTTCCCAATCTTCTATTATATGAT 180				
Db 121 tggcgaaacaagaagttgaattgggtttgaggtttcccaatcttctattatattgt 180				
0Y 181 GGTGATGTTAAATTAACACAGCTATAGGCCATCAATAGCTTATATGCTGACAAAC 240				
Db 181 ggtgatgtttaaattaacacagctataggccatcaatagcttataatagctgaaacaac 240				
0Y 241 AGCTTGCGTGTGTTGCCAAAGAGCGTGCAGAGATTTCAATCTTGAAGAGACGGTTTGG 300				
Db 241 agcttgcgtgtgttgccaaagagcggtgcagagatttcaatcttgaagagacggtttg 300				
0Y 301 GATTATTAGATACGGTGTTCGAGAAATGCAATTTAGTAAAGACTTTGAACCTCAAAATT 360				
Db 301 gattattagatacgggtgttcgagaaatgcgaatTTAGTAAAGACTTTGAACCTCAAAATT 360				
0Y 361 GATTTTCTTAGCAGCTACTCTGAATGCTGGAAGATTTGGAAGATGTTTATGTCATAA 420				
Db 361 gatTTTCTTAGCAGCTACTCTGAATGCTGGAAGATTTGGAAGATGTTTATGTCATAA 420				
0Y 421 ACATATTTAAATGCTGATCATGTAAACCATCTCGACTCATGTGTATGACGCTTTGAT 480				

Db	421	acatacttaaatgtagatcatgttaacccatccctgactcaatgltgtatgacgcctttat	480
Oy	481	GTGGTTTAAATACAGGACCCATATGCGCTGGATGGTTCCCAAAATATGTTTAA	540
Db	481	gtgtgttataacatgagcccaatgycgttgatgtgttccaaatagltgtttaa	540
Oy	541	AAACGATTTGAAGCTATCCACAAATGATTAAGTACTTGAATCCACAGATATATAGA	600
Db	541	aaacagtatgaagctatcccaaatatgataagtaacttgaatccagatataagca	600
Oy	601	TGGCCTTTGCAGGCGCTGCGAACCCACGTTTGGTGGTGCCACATCTCTCCAAATGGGAT	660
Db	601	tggcctttgcaagggctcgcgaacccaglttgglygttgcgcacatctccaaatcgat	660
Oy	661	CTGGTTCCGGGTGGATCCCGCAATTTCCGGGTCAGTCTGAACGCGCCGCTGATCAC	720
Db	661	ctggttccgggtgtagatccccggaattcccggtgtcgactccgagcgcgctgagatcac	720
Oy	721	AGGATCCCTCTGTACAAAGGCAAGTCTCTGAAGAGGCGCTGAAGAGCATGGGCTTCTG	780
Db	721	aggatccctctgtacaagaagcagltctctgaggaaagcgcgtagagagcatggctctg	780
Oy	781	GAGGACATCCCTGAGAGAAACAGAGATGGGCAATCGACAGACAGTATCCGGCTCGTCGC	840
Db	781	gaggacatccctgcagaagaacagcagatagtcacacgaagatctccggctctgcgttc	840
Oy	841	TATACGACTGTATCCGAGTCCGGTCGAGAACCTGTCTCTGTGAGGGTTCCAACTGTGT	900
Db	841	tatacgcactgtacgcagtcggtcgttcagaacctgtctctctgtgaggtttccaagctgt	900
Oy	901	GGTCAGGGTAAACAAGTGTATCTCGGTTTCCAGGCTGAGAGAACCATGTGTACCCGT	960
Db	901	ggtcagggttaacaagtagtactctcggttcgcacggtggaagaaccagtggtacccgt	960
Oy	961	GAGGAGACCCCAAGCGACAGTCTCCACAACGACGGTGTACTTGAAGAGATCCAGAGAG	1020
Db	961	gagggaaccccccaagccacagtcctccaacaagcagtgtagcttggagagatcccaagag	1020
Oy	1021	TATCTCAGTAAAGATCTAAGCTTGCTGCTGCTATCGAATTCGTGACGCCGGGGGATCC	1080
Db	1021	tatctccagtaaaagactaagcttgctgtgtctatcgaaattctgcagcccggtgagtc	1080
Oy	1081	ACTAGTTCTTAGACGG 1096	
Db	1081	actagttcttagagcgg 1096	
RESULT	2		
X01312			
ID	X01312	standard; cDNA to mRNA; 720 BP.	
XX	AC		
XX	X01312;		
XX			
XX	13-APR-1999 (first entry)		
XX			
DE	GST coding sequence.		
XX			
KW	Cell surface nonexpressive functional protein; cell-permeating protein		
KW	protein screening; protein production; GST; glutathione-S-transferase;		
OS	Schistosoma japonicum.		
XX			
XX	Jp11009278-A.		
XX			
XX	19-JAN-1999.		
XX			
XX	23-JUN-1997; 97JP-0165788.		
XX			
XX	23-JUN-1997; 97JP-0165788.		
XX			
PA	(SOME) SUMITOMO ELECTRIC IND CO		

XX WPI: 1999-145895/13.
 DR P-PSDB: W73909.
 XX
 PT Microbe having a nonexpressive functional protein on its surface -
 PT for screening and large-scale preparation of a selected function
 PT protein
 XX
 XX Disclosure; Page 9; 13pp; Japanese.
 XX
 CC This sequence encodes the glutathione-S-transferase (GST) protein, which
 CC was expressed in the microbe of the invention. The microbe has a cell
 CC surface nonexpressive functional protein on its surface, fused to a
 CC support protein derived from a cell-permeating protein. The microbe can
 CC be used for screening for biological activity of a cell surface
 CC nonexpressive protein. It can also be used for preparation of a selected
 CC functional protein in a large amounts.
 XX
 SO Sequence 720 BP; 209 A; 130 C; 164 G; 217 T; 0 other;

Query Match 62.5%; Score 685; DB 20; Length 720;
 Best Local Similarity 98.1%; Pred. No. 4.2e-179;
 Matches 705; Conservative 0; Mismatches 10; Indels 4; Gaps 1;

QY 1 ANGTCCTATCTAGCTTATTGGAATTAAGGCGCTTGCAACCCACGCTGCTT 60
 Db 1 atgtccctatactagctatgtgaaatlaagcgcttgcaaccactgactctt 60
 QY 61 TTGGATATCTTGACAAAAAATATGAAGACATTTGATGAGCGCGATGAGTGATAA 120
 Db 61 ttggaatactcttgaagaaaaataatgaagacatttgtatgaagcgatgaagtgataa 120
 QY 121 TGGCGAAGAAAAAGTTGAATGGTTGGAGTTGCCAATCTTCCTATTATTTGAT 180
 Db 121 tggcgaaacaaaagttgaatggttggagtttcccaatcttcctattatattgat 180
 QY 181 GGTGATGTTAATTAACACAGTCTATGCGCATCATGTTATATAGCTGACAGCAAC 240
 Db 181 ggtgatgtaataatgaacagctctatgcatcatagctatagctgcaagcaaac 240
 QY 241 ATGTTGGGTGGTTGTCACAAAGCGCTGACAGATTTGATGCTTAAGAGCGGTTTG 300
 Db 241 atgttgggtgggttgcacaaagcgctgacagatcttcaatgcttgaagcggtttg 300
 QY 301 GATATTAGATACGCTGTTTCGACAAATGCAATAGTAAAGACTTGAACCTGCACAA 360
 Db 301 gatattagatagctgttccgaaatgcaatagtaagacttgaacctgcacaaagt 360
 QY 361 GATTTTCTTAGCAGCTACCTGAAATGCTGAANAATGTTGAGAGTCTTATGCTATAA 420
 Db 361 gatttcttagcagctacctgaaatgctgaanaatgctgagatcgtttatgctataa 420
 QY 421 ACATATTAAAGTGATCATATGAACCATCCGACTTCATGTTGATGACGCTTTGAT 480
 Db 421 acatattaaagtgatcatatgaaccatccgacttcattgttlatgaagcgtttgat 480
 QY 481 GTTGTTTAACTGACCAATGCTGCTGATCGCTGCCAATAATAGTTTGTAA 540
 Db 481 gttgttttaactgacccaatgctgctgatcgctgcccaataatagtttgttaa 540
 QY 541 AAAGTATTGAAGCTATCCACAAATGATAGTACTTGAATCCAGCAAGTATATAGCA 600
 Db 541 aaagtattgaagctatccacaatgatagatgacttgaatccagcaagtatatagca 600
 QY 601 TGGCCTTTGACAGGCTGGCAGCAGCTTTGGTGCGACCATCTCCAAAATGGAT 660
 Db 601 tggcctttgacaggctggcagcagctttggtgcgacccatctccaaaatggat 660
 QY 661 CTGTTCCGCGT----GATCCCGCAATTCGCGGTCGATCGAGCGCGCCCTCGGA 715
 Db 661 ctgttcgaaggtcgtgtggtatcccgaaatcccggttcgactcgagcgccgcatcgga 715

RESULT 3
 ID 258406
 XX 258406 standard; cDNA; 3591 BP.
 AC 258406;
 XX
 DT 23-MAY-2000 (first entry)
 DE
 XX DNA encoding fumonisin esterase-amino polyol amine oxidase fusion.
 KW Amino polyol amine oxidase; APAO; fumonisin esterase; mycotoxin;
 KW transgenic plant; detoxification; animal feed; silage;
 KW glutathione S transferase; alpha-amylase; signal peptide; maize;
 SS.
 XX
 OS Chimeric - Schistosoma japonicum.
 OS Chimeric - Bacterium.
 OS Chimeric - Exophiala spinifera.
 XX
 FH Key Location/Qualifiers
 FT sig-peptide 1..687
 FT /tag- a
 FT /product- "GST - polylinker"
 FT mat-peptide 677..3588
 FT /tag- b
 FT /product- "fumonisin-esterase-K:trAPAO fusion"
 FT misc-feature 2200..2202
 FT /tag- c
 FT /note- "extra lysine codon"
 XX
 PN WO200004159-A1.
 XX
 PD 27-JAN-2000.
 XX
 PF 08-JUL-1999; 99WO-US15454.
 XX
 PR 15-JUL-1998; 98US-0092936.
 PR 21-MAY-1999; 99US-0135391.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 PA (CURA-) CURAGEN CORP.
 PI
 XX
 DR Duvick JP, Gilliam JT, Maddox JR;
 DR WPI: 2000-182425/16.
 DR P-PSDB; Y58917.
 XX
 PT New isolated polynucleotides, polypeptides useful for detecting and
 PT degrading fumonisin or structurally related mycotoxin in processed
 PT grain or in silage
 XX
 PS Example 13; Page 124-128; 154pp; English.
 XX
 CC The present sequence is that of a polynucleotide encoding a
 CC fusion protein (see Y58917) composed of glutathione S transferase
 CC and a bacterial fumonisin esterase (BST1) joined via a peptide
 CC linker to K:trAPAO (see Y58909), a truncated, but still functional,
 CC amino polyol amine oxidase of Exophiala spinifera. The construct
 CC was designed for expression in Escherichia coli of a hybrid
 CC protein having BST1 and APAO activity. The invention provides APAO
 CC polynucleotides (see 258383-87) and polypeptides (see Y58900-05) of
 CC E. spinifera and Rhinocladia atroviensis. The polynucleotides
 CC are used to transform plant cells normally susceptible to Fusarium
 CC or other toxin-producing fungus infection. Also provided are
 CC methods for expressing APAO, optionally as a fusion protein with
 CC fumonisin esterase, in transgenic plants, prokaryotic and non-plant
 CC eukaryotic systems. Methods for detoxification of grain, plant
 CC processing, silage, food crops and in animal feed and rumen
 CC microorganisms are also disclosed.
 XX
 SO Sequence 3591 BP; 758 A; 1017 C; 1064 G; 752 T; 0 other;

Query Match	62.48;	Score 684.2;	DB 21;	Length 3591;
Best Local Similarity	97.58;	Pred. No. 1.4e-178;		
Matches 695;	Conservative	0;	Mismatches 18;	Indels 0;
				Gaps 0;

QY	1	ATGTCCTCCATACAGAGTTATTTGGAAAATTAAGGGCGCTTGTGCAGACCCATCGACACTTCTT	60
Db	1	atgtcccccatactacagttatcttggaaaatttaaaggcccttctgaaacccactgcactctt	60
QY	61	TTGGAATATCTTTGAAGAAAAATATGAAGAGACATTTTGTATGAGCGCGATGNAAGGTGTATAA	120
Db	61	ttggatatcttcttgaaagaaaaatatgaagagcatcttgaatagcgcgatatgaaggtgtataa	120
QY	121	TGGCAACAAACAAAGTTTGAATGGGTGGTTGGAGTTTGGCAATCTCTCTTATTTATTTGAT	180
Db	121	tcggcaaacacaaagttctgaaatctgggttctggagttccccaactcccttatattatg	180
QY	181	GGTGATGTTAAATTAACACACATCTATGCGCATCATATACGTTATATATATATAGTGAACAGCACAC	240
Db	181	ggtgatgttaaatattcaacacagctctatggccatcctacgtatataagttgacaagcaaac	240
QY	241	ATGTTGGGTGGTGTGCCAAAAGACGCTGCGACAGATTTCAATGCTTGAAGAGCGGTTTG	300
Db	241	atgttggttggttcttccaaaagacgctgcagagtttcaatgtcttgaaggagcggtttg	300
QY	301	GATATTAGATACGGTGTTCGACAATTGCATATATGTAAAGACTTGTGAACCTGCAAGT	360
Db	301	gatatattagatacggctgtttctggagatctgcatalagtaagaactttgaacactccaaagt	360
QY	361	GATTTTCTTACCAAGCTACCTGGAATGCTGAAAAATGTTGCAAGACGTTTATGTCAATAA	420
Db	361	gatttctctagcaagctaacctgaaatgtctaaaatgttcgaagatcgttatgtcataaa	420
QY	421	ACATATTAAATGGGATGCATGTATACCCATCCGACATTCATGTTGTATATGACGGCTTGAT	480
Db	421	acataattaaatggtgatcatgatactgaaccacaccgaaactcgtctgtatagcgctcttgat	480
QY	481	GTTGTTTATATACATGAGCCCAATGCGCTGGATCGGTTCCCAAAATATAGTTTGTTTAAA	540
Db	481	gttcttattacacaggaacccaatgtgccttgatbgtctcccaaatagtgtgtttaaa	540
QY	541	AAAGCATTAAGAGCTATCCCAAAATGTATAGTACTTGAATTCGACAGAGTATATAGCA	600
Db	541	aaacgatatgaagcatatcccaataatgtataagtaacttgaataatccagcaagtatataga	600
QY	601	TGGCCTTTGCAGGGCTGAGCAAGCCAGCGTTTGAGTGCGACAGCATCTCCAAATATGAT	660
Db	601	tggcctttgcagggcttgcaagccaagctttgtgtgtggtgcagcatcctccaataatcgat	660
QY	661	CTGGTTCGCGCTGATCCCGCAATTCGCCGGGTGCATCTCAGACGCGCGCGCTGCA	713
Db	661	ctgtgttcgctgtgatacccggaatttcagagatlltccggttcgcagcagaccca	713
RESULT 4			
ID	Z60645		
XX	Z60645 standard; DNA; 3591 BP.		
AC	Z60645;		
XX	16-MAY-2000 (first entry)		
DT	DNA encoding a fusion of aminopolylol amine oxidase/fumonisin esterase		
DE			
XX	Aminopolylol amine oxidase: APAO; mycotoxin degradation; antimicrobial		
KW	plant fungal invasion; hydrogen peroxide; Fusarium; fungi;		
KW	funomisin esterase; ss.		
XX			
OS	Synthetic.		
OS	Exophiala spinifera.		
XX			
Key	Location/Qualifiers		
TH	1...3691		
CDS			

FT	/tag= a
FT	/product= "GST/aminopolylol amine oxidase/fumonisin
FT	esterase fusion"
FT	1..687
FT	/tag= b
FT	/note= "encodes GST and a linker"
FT	688..2163
FT	/tag= c
FT	/product= "bacterial funmonisin esterase"
FT	2164..2199
FT	/tag= d
FT	/note= "spacer sequence"
FT	2200..3588
FT	/tag= e
FT	/product= "aminopolylol amine oxidase"
FT	/note= "the protein has an amino terminal Lys
FT	for optimized expression"

PN	WC2000004160-AL.
XX	
PD	27-JAN-2000.
PF	08-JUL-1999; 99WC-US15455.
XX	
PR	15-JUL-1998; 98US-0092936.
XX	
PA	(PION-) PIONEER HI-BRED INT INC
PA	(CURA-) CORAGEN CORP.

XX	Crasta OR, Duvick J, Folkerts O, Gilliam JT, Maddox JR;
XX	WPI; 2000-182426/16.
DR	P-PsDB; Y68853.
XX	
PT	New nucleic acid encoding aminopolylol amine oxidase, used, e.g. to
PT	generate plants resistant to Fusarium -
XX	
PS	Example 13; Page 119-123; 145pp; English.
XX	
CC	The present sequence encodes a fusion protein of GST/Exophiala
CC	spluifera aminopolylol amine oxidase (APO)/Fumonisin esterase.
CC	The APO enzyme has homology to the flavin containing amine oxidase
CC	family, that oxidise primary amine to an aldehyde or ketone, releasing
CC	ammonia and hydrogen peroxide. The APO enzyme degrades mycotoxins
CC	that promote fungal invasion of plants. Destruction of mycotoxins by
CC	APO generates, as a by-product, hydrogen peroxide which is itself an
CC	antimicrobial and stimulates the plants own defensive systems. The
CC	APO polynucleotides are used to generate plants (particularly maize)
CC	that are resistant to Fusarium or other fungi that produce mycotoxins
CC	and/or to degrade such mycotoxins (e.g. during ensilaging); for
CC	recombinant production of APO polypeptides; as selection markers for
CC	plant transformation; and to isolate related sequences from other
CC	organisms. The APO polypeptides are used to degrade mycotoxins in
CC	plant materials, including expression in engineered bacteria and fungi
CC	e.g. rumen microflora.

Seq	Sequence	3591 BP;	758 A;	1017 C;	1064 G;	752 T;	0 other;
Query Match		62.4%;	Score 684.2;	DB 21;	Length 3591;		
Best Local Similarity		97.5%;	Prod. No. 1.4e-178;				
Matches	695;	Conservative	0;	Mismatches	18;	Indels	0;
						Gaps	
OY	1	ATGTCCTTACTACGTATTTGGAATAATTAAGGCGCTTGTGCACACCGACTCGACTTCTT	60				
Db	1	atgcccttactactaggttalttggaaattlaagggcctttgcacccactcgaactctt	60				
OY	61	TTGGAATATCTTCAGCAAAAATATGAGAGCATTTGTATGACGCCGCGATGAAGTGATPAA	120				
Db	61	ttggaatctcttgaagaanaataagaagcatttcatgaagcgcgatgaagtgataaa	120				
OY	121	TGCGGAACAAAAAGTTTGAATTGGCTTGGAGTTCACATCTTCCTTATTATTTGAT	180				


```

Db 121 tggcgaaacaaagtctgaattgggtttggagttcccaatcttccttattatattgat 180
Qy 181 GGTGATGTTAAATTAACACAGTCTATGAGCCATCATATGATATATAGCTGACACCAAC 240
    |||
Db 181 ggtgatgttaaatlaacacagctctatgacctatcatatagcttgacaaagcaaac 240
Qy 241 ATGTTGGGTGGTTCCTCAAAAGAGCTGCAGAGATTTCAATGCTTGAAGAGCGGTTTGG 300
    |||
Db 241 atgttggtggtgtgtcctcaaaagagcgtgcagagattcactgcttgaaagagcggttttg 300
Qy 301 GATATTAGTAGCGGTGTTTCGAGAAATGATATAGTAAGACTTTGAACCTCCAAAGTT 360
    |||
Db 301 gatattagatacgggtgttcgaaatcagatacagaagacttgaaactccaaagtt 360
Qy 361 GATTTCTTAGCAAGCTTACCTGTAATGCTGAATAATGTTGGAAGANTGTTTATGATCAATA 420
    |||
Db 361 gatttcttagcaagcttacctgaatgctgaataatgttcgaagatcgctttagtcaataa 420
Qy 421 ACATATTAAATGCGATCATGTACCCATCTGACTTCATGTTGATGACGCTTTGAT 480
    |||
Db 421 acatattaaatggtatcatgatgtaaccatccgacttcatgtgtatgacgtcttgat 480
Qy 481 GTTGTATTATACATGACCAATGTCCTGAGATGGTCCCAAAATAGTTTGTAA 540
    |||
Db 481 gtgttttatacactgagaccatgctgagcgttcccaaatagttgttttaaa 540
Qy 541 AAACGATTGAAGCTATCCACAAATGATAGTACTTGAATAATCCAGAAATGATATAGA 600
    |||
Db 541 aaacgattgaagctatcccaaatatgataagcttgaataatccgaagatataagca 600
Qy 601 TGGCGTTTGCAGGCGTGCAGACGCGTTTGTGTGCGGACCATCTCCAAATCGAAT 660
    |||
Db 601 tggcgttgcagggcgagcgaagcgcgttgggtggtgacacatccctcaaatcgat 660
Qy 661 CTGGTCCGCGGATGATCCCGAATTCGCGGTGACTCGACGCGCGCGTGA 713
    |||
Db 661 ctggttcggtggtggtatcccggaattcgcagatcttcggtccgcagagccga 713

```

RESULT 5

232916
ID Z32916 standard; CDNA: 2475 BP.

AC Z32916;

DT 09-FEB-2000 (first entry)

DE Human HCP/GST fusion gene.

KW Haematopoietic cell phosphatase; HCP; cloning; expression;

KW epithelial cell; growth factor receptor; interleukin-3; IL-3;

KW erythropoietin; EPO; negative regulator; signal transduction;

KW modulation; activity; inhibitor; stimulation; cytokine therapy;

KW antisense therapy; gene therapy; treatment; diagnosis;

KW glutathione-S-transferase; fusion; ds.

OS Synthetic.

OS Homo sapiens.

PH Key location/Qualifiers

FT CDS 1..2475 /tag= a

FT /product= "Human HCP/GST fusion protein"

PN WO9954450-A1.

PD 28-OCT-1999.

PF 15-APR-1999; 99WO-US08228.

PR 17-APR-1998; 98US-0082072.

XX (ORTH) ORTHO-MCNEIL PHARM INC.

```

XX Jolliffe L, Barbone F, Shanahan M, Xu D;
PI MPI: 2000-013244/01.
DR F-PSDB; Y52288.
XX Assay methods for modulators of haematopoietic cell phosphatase -
PT Example 3; Fig 4A; 57Pp; English.
PS
XX This sequence represents a full-length human haematopoietic
CC cell phosphatase (HCP)/glutathione-S-transferase (GST)
CC fusion gene. This was cloned and expressed in E. coli and
CC subsequently purified for use in a variety of activity assays.
CC HCP is a 68 kD non-membrane associated protein found primarily in
CC haematopoietic cells, and to some extent in epithelial cells. It
CC has been shown to associate with several growth factor receptors
CC following ligand stimulation and acts as a negative regulator
CC of receptor activation, and hence of haematopoietic signal transduction.
CC The receptors it regulates include the interleukin-3 (IL-3) and
CC erythropoietin (EPO) receptors. The invention relates to a novel
CC method of identifying compounds that modulate HCP activity. This
CC comprises combining a modulator of HCP protein activity with HCP protein
CC and HCP protein substrate, and measuring an effect of the modulator on
CC the activity of the HCP protein. Inhibitors of HCP can be used to
CC modulate the activity of haematopoietic receptors. HCP has been shown to
CC be a negative regulator of one or more signal transduction pathways in
CC haematopoietic cells, and the identification of a HCP inhibitor would
CC provide a synthetic stimulator to the haematopoietic system could be
CC used in conjunction with other cytokine therapy, e.g., administration of
CC EPO. HCP is also associated with several growth factors found in
CC haematopoietic cells and it is possible that a HCP inhibitor may also
CC function as a haematopoietic cell potentiator, enhancing the effect of
CC growth factors, decreasing the necessary dose of the factor. The HCP
CC polynucleotides are sources of probes and primers, and can be used to
CC design antisense sequences, and in gene therapy. The protein, or
CC active fragments thereof, may be used to treat or diagnose disorders
CC which require the modulation of the HCP activity. Small molecules that
CC inhibit the action of HCP can also augment the effect of EPO, which
CC meaning that a lower therapeutic dose of EPO may be utilised.
XX
SQ Sequence 2475 BP; 638 A; 612 C; 701 G; 524 T; 0 other;

```

Query Match 62.4%; Score 684; DB 21; Length 2475;
Best Local Similarity 99.3%; Pred. No. 1.4e-178;
Matches 667; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

Qy 1 ATGTCCCTATAGTATGGAATAATTAAGGCGCTTGTGCAACCACTGCACTTCTT 60
    |||
Db 1 atgtccctatactagtagtataaataaagcgcttgtaaacacactgcactctt 60
Qy 61 TTGGAATATCTTGAAGAAATATGAAGAGCATTTGTATGAGCGCGATGAGGTGATAA 120
    |||
Db 61 ttggaatattcttgaaagaaataatgaagcatttgatgagcgatgaaggtgataaa 120
Qy 121 TGGCGAACAACAAAGTTGATGGGTTGGAGTTTCCCAATCTTCTTATTTATGAT 180
    |||
Db 121 tggcgaaacaaagtttgatgggttggagtttcccaatcttccctattatattgat 180
Qy 181 GGTGATGTTAAATTAACACAGTCTATGAGCCATCATATGATATATAGCTGACACCAAC 240
    |||
Db 181 ggtgatgttaaatlaacacagctctatgacctatcatatagcttgacaaagcaaac 240
Qy 241 ATGTTGGGTGGTTCCTCAAAAGAGCTGCAGAGATTTCAATGCTTGAAGAGCGGTTTGG 300
    |||
Db 241 atgttggtggtgtgtcctcaaaagagcgtgcagagattcactgcttgaaagagcggtttg 300
Qy 301 GATATTAGTAGCGGTGTTTCGAGAAATGATATAGTAAGACTTTGAACCTCCAAAGTT 360
    |||
Db 301 gatattagatacgggtgttcgaaatcagatacagaagacttgaaactccaaagtt 360
Qy 361 GATTTCTTAGCAAGCTTACCTGTAATGCTGAATAATGTTGGAAGANTGTTTATGATCAATA 420

```

```

Db      361 gatttcttagcaagcaccgtaaatgctgaataatgctgaagatcgtttatgtaataa 420
Oy      421 ACATATTAAATGSGTGCATGTAACCCATCCGACTTTCATGTTGATGACGCTTGAT 480
Db      421 acataatlaaatgctgtaacatgtaacccatccctactcaltgtatgtaagcgtctg 480
Oy      481 GTTGTTTATACATGAGCCCAATGTGCTGATGCTTCCCAAAATAGTTGTTTAA 540
Db      481 gttctttatacaaggacccaatgctgtagagcgttcccaaatagtttttaa 540
Oy      541 AACGTTTAAAGTATCCCAAAATTTGATTAAGTACTGAAATCCAGCAATATATAGCA 600
Db      541 aaacgatttgaagctatcccaaatgtaagtaacttgaatccagcaagatatagca 600
Oy      601 TGGCCTTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
Db      601 tggcctttgagagcgtgagcagcagcagcagcagcagcagcagcagcagcagcagcag 660
Oy      661 CTGGTTCCGCGTGAGTCCCGAATTCCTCCGGGT 692
Db      661 ctggtccgctgtgattcccggaattcatggt 692

```

RESULT 6

Z32917 232917 standard; cDNA; 1866 BP.

AC Z32917;

DT 09-FEB-2000 (first entry)

DE Human HCP phosphodomain/GST fusion gene.

XX Haematopoietic cell phosphatase; HCP; cloning; expression;

KW epithelial cell; growth factor receptor; interleukin-3; IL-3;

KW erythropoietin; EPO; negative regulator; signal transduction;

KW modulation; activity; inhibitor; stimulation; cytokine therapy;

KW antisense therapy; gene therapy; treatment; diagnosis; phosphodomain;

KW catalytic domain; glutathione-S-transferase; fusion; ds.

XX Synthetic.

OS Homo sapiens.

XX Key

PF 15-APR-1999; 99WO-US08228.

PR 17-APR-1998; 98US-0082072.

PA (ORTH) ORTHO-MCMEIL PHARM INC.

PI Jolliffe L, Barbone F, Shanahan M, Xu D;

DR WPI; 2000-013244/01.

XX P-PSDB; Y52289.

XX Assay methods for modulators of hematopoietic cell phosphatase

XX Example 2; Fig 4C; 57pp; English.

CC This sequence represents a human haematopoietic cell

CC phosphatase (HCP) catalytic domain/glutathione-S-transferase (GST)

CC fusion gene. This was cloned and expressed in E. coli and

CC subsequently purified for use in a variety of activity assays.

CC HCP is a 68 kD non-membrane associated protein found primarily in

```

CC haematopoietic cells, and to some extent in epithelial cells. It
CC has been shown to associate with several growth factor receptors
CC following ligand stimulation and acts as a negative regulator
CC of receptor activation, and hence of haematopoietic signal transduction.
CC The receptors it regulates include the interleukin-3 (IL-3) and
CC erythropoietin (EPO) receptors. The invention relates to a novel
CC method of identifying compounds that modulate HCP activity. This
CC comprises combining a modulator of HCP protein activity with HCP protein
CC and HCP protein substrate, and measuring an effect of the modulator on
CC the activity of the HCP protein. Inhibitors of HCP can be used to
CC modulate the activity of haematopoietic receptors. HCP has been shown to
CC be a negative regulator of one or more signal transduction pathways in
CC haematopoietic cells, and the identification of a HCP inhibitor would
CC provide a synthetic stimulator to the haematopoietic system could be
CC used in conjunction with other cytokine therapy, e.g., administration of
CC EPO. HCP is also associated with several growth factors found in
CC haematopoietic cells and it is possible that a HCP inhibitor may also
CC function as a haematopoietic cell potentiator, enhancing the effect of
CC growth factors, decreasing the necessary dose of the factor. The HCP
CC polynucleotides are sources of probes and primers, and can be used to
CC design antisense sequences, and in gene therapy. The protein, or
CC active fragments thereof, may be used to treat or diagnose disorders
CC which require the modulation of the HCP activity. Small molecules that
CC inhibit the action of HCP can also augment the effect of EPO, which
CC meaning that a lower therapeutic dose of EPO may be utilised.
CC
XX
SO

```

Sequence 1866 BP; 515 A; 444 C; 503 G; 404 T; 0 other;

Query Match 62.4%; Score 683.4; DB 21; Length 1866;
Best Local Similarity 99.1%; Pred. No. 1.8e-178;
Matches 687; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

Oy      1 ATGTCCCTATACGTTATTTGCAAAATTAAGGCGCTTGACACCCACTGACTTCTT 60
Db      1 atgtccctatacagttatcttgaataatlaaggccttgcacaccacgcgtctctt 60
Oy      61 TTGCAATATCTTGAAGAAATATGAAGAGCATTTGATGAGCGCGATGAAAGTGAT 120
Db      61 ttgcaatctctgaagaaataatgaagagcatcttgatgagcgatgaaggatgataa 120
Oy      121 TGGCGAACAANAAGTTGAATGGGTTTGAGATTGCCAATCTCCCTATATATGAT 180
Db      121 tggcgaaacaaaagttgaattgggtttgagattgccaatctccctatatatgat 180
Oy      181 GGTGATGTTAAATTAACAGCTCATGCGCATCTATAGTTATATAGTGCAAGCACAC 240
Db      181 ggtgatgtaaatttaacagctcatgcgcatctatagttatagtgcaagcacac 240
Oy      241 ATGTGGGTGGTTGTCCAAAAGCGTGACAGATTTCATGCTTGAAGAGGCGGTTTG 300
Db      241 atgtgggtggtgtgtccaaaagcgtgacagatttcattgcttgaaggcggtttg 300
Oy      301 GATATTAAGATAGCGTGTGTCGAGATTTGATTAAGAAAGCTTGAACCTCAAGTT 360
Db      301 gatattagatagcgtgtgttcgagattgcatatagtaagaagcttgaacctccaagt 360
Oy      361 GATTTCTTACAGACTACCTGGAATGCTGAATGTTGCAAGATGTTATATGCTATAA 420
Db      361 gatttctttagcagactacctgaatgctgaatgctgaatgctgaatgctgaatgctgaat 420
Oy      421 ACATATTAAATGSGTGCATGTAACCCATCCGACTTTCATGTTGATGACGCTTGAT 480
Db      421 acataatlaaatgctgtaacatgtaacccatccctactcaltgtatgtaagcgtctg 480
Oy      481 GTTGTTTATACATGAGCCCAATGTGCTGATGCTTCCCAAAATAGTTGTTTAA 540
Db      481 gttgtttatacaaggacccaatgctgtagagcgttcccaaatagtttttaa 540
Oy      541 AACGTTTAAAGTATCCCAAAATTTGATTAAGTACTGAAATCCAGCAATATATAGCA 600
Db      541 aaacgatttgaagcgtatcccaaatgtaagtaacttgaatccagcaagatatagca 600

```

```
OY 601 TGGCCTTTGACGGCTGGACACGACGTTGGTGGTGGGACACCTCCAAATGCGAT 660
    |||
DB 601 Tggccttgcagagctgcgaagccacgcttgcgtgcgcaccatctccaaatcgat 660
OY 661 CTGGTTCGCGGTGATCCCGGAATTCGCCGGGTC 693
    |||
DB 661 ctggttcgcgctggtatcccggaattcgcgcc 693

RESULT 7
ID 258400 standard; cDNA; 2079 BP.
XX 258400;
AC 258400;
DT 23-MAY-2000 (first entry)
XX
DE DNA encoding GST-truncated amino polyol amine oxidase fusion.
XX
KW Amino polyol amine oxidase; APAO; fumonisin; mycotoxin;
KM transgenic plant; detoxification; animal feed; silage;
XX selectable marker; glutathione S transferase; GST; ss.
OS Chimeric - Schistosoma japonicum.
XX Chimeric - Exophiala spinifera.
XX
FH Key Location/Qualifiers
FT sig_peptide 1..687
    /tag= a
    /product= "gst fusion + polylinker"
FT mat_peptide 688..2076
    /tag= b
    /product= "K:trPAO"
FT misc_feature 688..690
    /*tag= c
    /note= "extra lysine"
FT
XX
XX WO200004159-A1.
XX
XX 27-JAN-2000.
XX
XX 08-JUL-1999; 99WO-US15454.
XX
XX 15-JUL-1998; 98US-0092936.
XX
XX 21-MAY-1999; 99US-0135391.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX (CURA-) CURAGEN CORP.
XX
XX Duvalck JP, Gilliam JT, Maddox JR;
XX
XX WPI: 2000-182425/16.
XX
XX P-PSDB; Y58911.
XX
XX New isolated polynucleotides, polypeptides useful for detecting and
XX degrading fumonisin or structurally related mycotoxin in processed
XX grain or in silage -
XX
XX Example 9; Page 89-92; 154pp; English.
XX
XX The present sequence is that of a polynucleotide encoding a
XX GST:K:trPAO fusion protein composed of glutathione S transferase
XX and K:trPAO, a truncated, but functional, amino polyol amine
XX oxidase (see also Y58909) of Exophiala spinifera. The construct
XX was designed for expression of K:trPAO in Escherichia coli.
XX K:trPAO is capable of degrading fumonisin and related mycotoxins.
XX The invention provides APAO polynucleotides (see Z58383-B7) and
XX polypeptides (see Y58900-05) of E. spinifera and Rhinocladiella
XX atrovirens. The polynucleotides are used to transform plant cells
XX normally susceptible to Fusarium or other toxin-producing fungus
XX infection. Also provided are methods for expressing APAO in
XX transgenic plants, prokaryotic and non-plant eukaryotic systems.
XX Methods for detoxification of grain, grain processing, silage, food
```

```
CC crops and in animal feed and rumen microorganisms are also disclosed.
CC APAO polynucleotide can also be used as a selectable marker.
XX
XX Sequence 2079 BP; 537 A; 490 C; 547 G; 505 T; 0 other;
SQ
Query Match 62.3%; Score 683.2; DB 21; Length 2079;
Best Local Similarity 98.9%; Pred. No. 2.1e-178;
Matches 688; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
OY 1 ATGTCCCTTACTAGCTATTTGAAATTAAGGCGCTGTGCAACCCACTGACTTCTT 60
    |||
DB 1 atgtcccttactagctatlttggaaattaaaggccttgcgaaccactcgacttctt 60
OY 61 TTGAATATCTTGAAGAAATATGAGAGCATTTGTATGAGCGCGATGAGTGATTA 120
    |||
DB 61 ttggaatattcttgaagaaatataatgaagcattgtatgagcgcgatgaagtataa 120
OY 121 TGGCGAAMCAAAAAGTTGAATTTGGATTGCCAATCTTCTATTATATTTGAT 180
    |||
DB 121 tggcgaaacaaaagttgaattggattggattgcccaatcttctattatattgt 180
OY 181 GGTGATGTTAAATTAACACAGCTATAGCCATCATACGTTATATACCTGACACAC 240
    |||
DB 181 ggtgatgttaaattaacacagctatagccatcatacgttatatactgacaaagcaac 240
OY 241 ATGTTGGTGGTGTGTCGCAAAAGAGCGTGACAGATTTCATGCTTGATGAGGCGTTTG 300
    |||
DB 241 atgttggtggtgtgtccaaaagagcgctgcagagatttcattgctgaagagcgcttg 300
OY 301 GATATTAGATACGCTGTTTCGAGAAATGATATAGTAAGACTTTGAAACTCTCAAA 360
    |||
DB 301 gatattagatacgctgtttcgagaaatgatatagtaaagactttgaaactctcaaa 360
OY 361 GATTTTCTTAGCAAGCTACTGAAATGCTGAAAAATTTTGAGACATGTTATGTCTAA 420
    |||
DB 361 gattttcttagcaagctactgaaatgctgaaaaattttgagacatgtttatgtctaa 420
OY 421 ACATATTTAAATGATGATCATGTAACCCATCCGACTTCATGTTATGAGCGCTTGAT 480
    |||
DB 421 acatatttaaatgatatcatgttaacccatccgacttcattgattgagcgcttgat 480
OY 481 GTTGTTTTATACATGACCAACCAATGTGCTGATGCTGTCCTCCAAATATGTTGTTTAA 540
    |||
DB 481 gttgttttatacatgacccaaccaaattgtgctgatgctgctccaaatattgttttaa 540
OY 541 AAACGATTGAAGCTATCCCAAAATTGATTAAGTACTTGAATCCAGCAAGTATATGA 600
    |||
DB 541 aaacgattgaagctatcccaaaattgattaaagtacttgaatccagcaagtatalaga 600
OY 601 TGGCCTTTGACGGCTGGACACGACGTTGGTGGTGGGACACCTCCAAATGCGAT 660
    |||
DB 601 Tggccttgcagagctgcgaagccacgcttgcgtgcgcaccatctccaaatcgat 660
OY 661 CTGGTTCGCGGTGATCCCGGAATTCGCCGGGTCGAC 696
    |||
DB 661 ctggttcgcgctggtatcccggaattcgcgcc 696

RESULT 8
ID 260639 standard; DNA; 2079 BP.
XX 260639;
AC 260639;
DT 16-MAY-2000 (first entry)
XX
DE DNA encoding an aminopolyol amine oxidase/GST fusion.
XX
KW Aminopolyol amine oxidase; APAO; mycotoxin degradation; antimicrobial;
XX plant fungal invasion; hydrogen peroxide; Fusarium; fungi; ss.
XX
XX Synthetic.
```

```

OS Exophiala spinifera.
XX
FH Key Location/Qualifiers
FT CDS 1..2079
FT /tag= a
FT /product= "aminopolylol amine oxidase/GST fusion"
FT misc_feature 1..687
FT /tag= b
FT /note= "encodes GST and a linker"
FT mat_peptide 688..2076
FT /tag= c
FT /note= "aminopolylol amine oxidase, with an extra
FT amino terminal lys residue for optimized
FT expression"
XX
XX WO200004160-A1.
XX
XX 27-JAN-2000.
XX
XX 08-JUL-1999; 99WO-US15455.
XX
XX 15-JUL-1998; 98US-0092936.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX (CURA-) CURAGEN CORP.
XX
XX Crasta OR, Duvick J, Folkerts O, Gilliam JT, Maddox JR;
XX WPI: 2000-182426/16.
XX DR P-PSDB; Y68847.
XX
XX New nucleic acid encoding aminopolylol amine oxidase, used, e.g. to
XX generate plants resistant to Fusarium
XX
XX Example 9; Page 90-93; 145pp; English.
XX
XX The present sequence encodes a Exophiala spinifera aminopolylol amine
XX oxidase (APAO)/GST fusion. The APAO enzyme has homology to the flavin
XX containing amine oxidase family, that oxidize primary amine to an
XX aldehyde or ketone, releasing ammonia and hydrogen peroxide. The APAO
XX enzyme degrades mycotoxins that promote fungal invasion of plants.
XX Destruction of mycotoxins by APAO generates, as a by-product, hydrogen
XX peroxide which is itself an antimicrobial and stimulates the plants
XX own defensive systems. The APAO polynucleotides are used to generate
XX plants (particularly maize) that are resistant to Fusarium or other fungi
XX that produce mycotoxins and/or to degrade such mycotoxins (e.g. during
XX ensiling); for recombinant production of APAO polypeptides; as
XX selection markers for plant transformation; and to isolate related
XX sequences from other organisms. The APAO polypeptides are used to
XX degrade mycotoxins in plant materials, including expression in
XX engineered bacteria and fungi, e.g. rumen microflora.
XX
XX Sequence 2079 BP; 537 A; 490 C; 547 G; 505 T; 0 other;
XX
XX
XX Query Match 62.3%; Score 683.2; DB 21; Length 2079;
XX Best Local Similarity 98.9%; Pred. No. 2,1e-178;
XX Matches 688; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
XX
XX 1 AATGCCCTATAGTGGTATTTGAAATTAAGGCGCTTGCGAACCCAGCTGCTT 60
XX |
XX 1 atgtccctatacaggtatctggaataatgaagccttggaacccagctgactctt 60
XX
XX 61 TTGGAATATCTTGAAGAAAATATGAGAGCATTTGTATGAGCGATGAAGTGATAAA 120
XX |
XX 61 ttggaataatcttgaagaaaatatagaagcatttgatagcgcgatgaagtgataaa 120
XX
XX 121 TGGGAAACAAACAAATTTGAATTTGGGTTTGGAGTTTCCCAATCTTCTATTATTTGAT 180
XX |
XX 121 tgggaaacaaacaaatgttgaatttgggttggagtttcccaatcttctattattgat 180
XX
XX 181 GGTGATGTTAAATTAACACAGCTATGCGGCTCATGCTATATATAGCGAACACACAC 240
XX |
XX 181 ggtgatgtttaaattaacacagctatgcgccctcatgctatattatagcgaaacacac 240

```

```

Db 181 ggtgatgtttaaataacacagctatgcgccatcatgaagttatatagctgacaaacacac 240
Oy 241 ATGTGGGTGTTGTCCTCAAAAGAGCGTCAGAGATTTCAATGCTGAAGAGCGGTTTGG 300
|
|
|
Db 241 atgttggtgttgcctcaaaagagcgtcgaagatttcaatgctgaaagcggttttg 300
Oy 301 GATATTAGATACGGTGTTCGAGATTCATATACTATAAGACTTTGAACACTCTCAAGTT 360
|
|
|
Db 301 gattttagatagcgtgttgcgaagattgcatataglaaagacttgaacacttcaagtt 360
Oy 361 GATTTTCTTAGCAGCTACCTGTAATGCTGAAATGTTTCGAAGATCTTATGTCATAAA 420
|
|
|
Db 361 gatttcttagaagactacccgaagtgtgaatgttgcgaagatcgattatgcatataa 420
Oy 421 ACATTTTAAATGATGATCATATGTAACCATCTGACTTCATGTTGATGACGCTCTTGAT 480
|
|
|
Db 421 acattttaaatgtagatcatatgtaacccatctcatctcatgattgttagaagccttcat 480
Oy 481 GTTGTATTATACATGAGACCCCAATGCGCTGATGCTGCCAAATTAAGTTGTTTAAA 540
|
|
|
Db 481 gtgttttatacaagagaccacaatgacctgagatgcgttcccaaatltagttgttttaa 540
Oy 541 AACGTTATGAGCTATCCACAAATTTGATTAAGTACTTGAATTCAGCAAGTATATAGCA 600
|
|
|
Db 541 aaagttatgaagctatcccaacaattgataagtaacttgaatccagcaagtatataagca 600
Oy 601 TGGCTTTGCAAGGCTGCGCAAGCCAGCTTTGTGTGTGCGACATCTCCAAATTCGGAT 660
|
|
|
Db 601 tggcctttgcagggctgcgcaagccagctttgtgtgtgcgacatctccaaatcgat 660
Oy 661 CTGCTCCGCGTGGATGCGCCGCAATTCCTCGGCTGCAC 696
|
|
|
Db 661 ctgctccgcgtgtagtcccggaattcaagaacacac 696

```

RESULT 9

258407
ID 258407 standard; cDNA; 2490 BP.

XX AC 258407;

XX 23-MAY-2000 (first entry)

XX DNA encoding GST-truncated amino polylol amine oxidase fusion.

XX KW Amino polylol amine oxidase; APAO; fumonisin; mycotoxin;

XX KW transgenic plant; detoxification; animal feed; silage;

XX KW selectable marker; glutathione S transferase; GST; mutant; ss.

XX OS Chimeric - Schistosoma japonicum.

XX OS Chimeric - Exophiala spinifera.

XX OS Synthetic.

XX

FH Key Location/Qualifiers

FH sig_peptide 1..687

FT /tag= a

FT /product= "gst fusion + polylolinker"

FT mat_peptide 688..2076

FT /tag= b

FT /product= "K:trAPAO"

FT misc_feature 688..690

FT /tag= c

FT /note= "extra lysine"

FT mutation replace(1288..1290, TCC)

FT /tag= d

FT mutation replace(1303..1305, AAC)

FT /tag= e

XX

XX WO200004159-A1.

XX

XX 27-JAN-2000.

XX

XX 08-JUL-1999; 99WO-US15454.

XX 15-JUL-1998; 98US-0092936.
 PR 21-MAY-1999; 99US-0135391.
 XX (PION-) PIONEER HI-BRED INT INC.
 PA (CURA-) CURAGEN CORP.
 XX DuVick JP, Gilliam JT, Maddox JR;
 PI WPI; 2000-182425/16.
 DR P-PSDB; Y58918.
 XX New isolated polynucleotides, polypeptides useful for detecting and
 PT degrading fumonisin or structurally related mycotoxin in processed
 PT grain or in silage -
 XX Example 15; Page 131-135; 154pp; English.
 PS The present sequence is that of a polynucleotide encoding a
 CC GST:APAO fusion protein (see Y58918) composed of glutathione S
 CC transferase and Exophiala spinifera amino polyol amine oxidase
 CC (APAO). 2 Codons of the APAO coding region have been altered by
 CC site-directed mutagenesis in order to change the glycosylation
 CC pattern of the fusion protein. These were the replacement of AAT
 CC (Asn) by TCC (Ser) at codon 430, and replacement of AGC (Ser) by
 CC AAC (Asn) at codon 435. APAO is capable of degrading fumonisin and
 CC related mycotoxins. The invention provides APAO polynucleotides
 CC (see Z58383-87) and polypeptides (see Y58900-05) of E. spinifera
 CC and Rhinocladia atrovirens. The polynucleotides are used to
 CC transform plant cells normally susceptible to Fusarium or other
 CC toxin-producing fungus infection. Also provided are methods for
 CC expressing APAO in transgenic plants, prokaryotic and non-plant
 CC eukaryotic systems. Methods for detoxification of grain, grain
 CC processing, silage, food crops and in animal feed and rumen
 CC microorganisms are also disclosed.
 XX Sequence 2490 BP; 625 A; 623 C; 653 G; 589 T; 0 other:
 SO

Query Match 62.3%; Score 683; DB 21; Length 2490;
 Best Local Similarity 99.3%; Pred. No. 2.6e-178;
 Matches 686; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

0Y 1 ATGTCCTATATAGTATTGAAATTAAGGCGCTTGCAACCCACTGACTTCTT 60
 DB 1 atgtccctatactagtattgaaataaagagccttgcaaccactgacttctt 60
 0Y 61 TTGGAATATCTTGGAAGAAAATATGAGAGCATTTGTATGACCGGATGAGTATAA 120
 DB 61 ttggaatatcttggaagaaaataatgaagacattgtatgaagcgcgaltgaataa 120
 0Y 121 TGGCAACAAAGATTGCAATTTGGTTGGATTGCCAATCTTCCTATTATATTGAT 180
 DB 121 tggcaacaaagattgcaatttgggttggattggagttcccaacttcctattatattgat 180
 0Y 181 GGTGATGTTAAATTAACACAGTCTATGGCCATCATACGTTATATAGCTGACAGCAAC 240
 DB 181 ggtgatgtttaaattaacacagctctatggccatcatacgttatatagtcgaagaagc 240
 0Y 241 ATGTTGGTGGTGTTCCTCAAAAGAGCGTGACAGATTTCATATCTGAGGAGCGGTTTG 300
 DB 241 atgttgggtgtgttcctcaaaagagcgtgacagatttcattatcttgaagagcg 300
 0Y 301 CATATTAGATACGGTGTTCGAGAAATGCAATATAGTAAAGACTTTGAACCTGCAAGTT 360
 DB 301 gatattagatacgggtgttcgagaaatgcattatagtaaaagacttgaacacttcaagtt 360
 0Y 361 GATTTTCTTAGACACTACCTGAATGCTGAAGATGTTGGAAGATCGTTATGTCATATA 420
 DB 361 gatcttcttagacactacctgaatgctgaagatgtttggaagatcgttatgtcataa 420
 0Y 421 ACATATTTAAATGCGATCATGTATACCCATCTGACTTCATTTGTATAGCGCTTGTAT 480
 DB 421 acatattttaaattgcgatcatgtataccatctgacttcattttgtatagcgcttgtat 480

DB 421 acatattttaaattgctgcatgtataccatctgacttcattgtatgacgcttcat 480
 0Y 481 GTTGTATTATACATGACCAATGTCGTGATGCGTCCCAAAATTAAGTTGTTTAA 540
 DB 481 gtgtgtattatatacatgacccaatgctgctggaagcgttcccaaatatgttcttaa 540
 0Y 541 AAACGATTGAAGCTATCCACAAATTAAGTAAATGATTTGAAATCCAGCATATATAGCA 600
 DB 541 aaacgattgaagctatcccaacaattgaatgaatgaaatccagcaagtatatagca 600
 0Y 601 TGGCCTTTGACAGGCTGCGAAGCCAGCTTTGCTGTGGCAGCATCTCCAAATTCGAT 660
 DB 601 tggcctttgacagcgtgcggaagccagctttgctgtggcagcatctccaaatcga 660
 0Y 661 CTGCTTCGCGGTGATCCCGAATTCCCGG 691
 DB 661 ctgcttcgcggtgatcccggaattcccggaattcctcga 691

RESULT 10
 ID Z60646
 XX Z60646 standard; DNA: 2490 BP.
 AC Z60646;
 DR 16-MAY-2000 (first entry)
 XX
 DE DNA encoding a fusion of GST/mutated aminopolylol amine oxidase.
 XX
 KW Aminopolylol amine oxidase; APAO; mycotoxin degradation; antimicrobial;
 KW plant fungal invasion; hydrogen peroxide; Fusarium; fungi; ss.
 XX
 OS Synthetic.
 XX Exophiala spinifera.
 XX
 FT Key Location/Qualifiers
 FT CDS 1..2490
 FT /tag= a
 FT /product= "GST/aminopolylol amine oxidase fusion"
 FT misc-feature 1..687
 FT /tag= b
 FT /note= "encodes GST and a linker"
 FT mat-peptide 688..2487
 FT /tag= c
 FT /product= "aminopolylol amine oxidase"
 FT /note= "the protein has been mutated to remove potential glycosylation sites at positions 201 and 206"
 FT
 XX
 XX WO200004160-A1.
 PD 27-JAN-2000.
 XX 08-JUL-1999; 99WO-US15455.
 PR 15-JUL-1998; 98US-0092936.
 XX (PION-) PIONEER HI-BRED INT INC.
 PA (CURA-) CURAGEN CORP.
 XX Crasta OR, DuVick J, Folkerts O, Gilliam JT, Maddox JR;
 PI WPI; 2000-182426/16.
 DR P-PSDB; Y68854.
 XX New nucleic acid encoding aminopolylol amine oxidase, used, e.g. to
 PT generate plants resistant to Fusarium -
 XX Example 15; Page 127-130; 145pp; English.
 CC The present sequence encodes a fusion protein of GST and a mutated
 CC Exophiala spinifera aminopolylol amine oxidase (APAO). The APAO
 CC protein has been mutated to remove potential glycosylation sites at

CC positions 201 and 206. The APO enzyme has homology to the flavin
 CC containing amine oxidase family, that oxidise primary amine to an
 CC aldehyde or ketone, releasing ammonia and hydrogen peroxide. The APO
 CC enzyme degrades mycotoxins that promote fungal invasion of plants.
 CC Destruction of mycotoxins by APO generates, as a by-product, hydrogen
 CC peroxide which is itself an antimicrobial and stimulates the plants
 CC own defensive systems. The APO polynucleotides are used to generate
 CC plants (particularly maize) that are resistant to Fusarium or other
 CC fungi that produce mycotoxins and/or to degrade such mycotoxins
 CC (e.g. during ensiling); for recombinant production of APO
 CC polypeptides; as selection markers for plant transformation; and to
 CC isolate related sequences from other organisms. The APO polypeptides
 CC are used to degrade mycotoxins in plant materials, including expression
 CC in engineered bacteria and fungi, e.g. rumen microflora.
 SQ Sequence 2490 BP; 625 A; 623 C; 653 G; 589 T; 0 other;

Query Match 62.3%; Score 683; DB 21; Length 2490;
 Best Local Similarity 99.3%; Pred. No. 2.6e-178;
 Matches 686; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

1 ATGCCCCCTACTAGCTTATGGAAATTAAGGCGCTGTGCAACCCGACTGCTTCT 60
 1 atgccccctactagcttattggaaattaaaggcgctgtgcaacccgactgcttct 60
 61 TTGCAATATCTTGAAGAAAATATGAAGAGCATTTGTATGAGCGGATGAAGTATAA 120
 61 ttgcaatatcttgaagaaaatatagaagagcattgtatgagcggatgaagtataa 120
 121 TGGCGAAGAAAAGTTTGAATGGGTTTGGAGTTTCCCAATCTTCTTATTAATTGAT 180
 121 tggcgaagaaaagtttgaatgggtttggagtttcccaatcttcttattatgattg 180
 181 GGTGATGTTAATTAATTAACAGCTATGGCATCATGCTTATATAGTGACAGACAGAC 240
 181 ggtgatgtttaatttaataacagctatggcatcatgcttattatagtgacagacag 240
 241 ATGTGGGTGGTGTCCAAAAGAGCGTGACAGATTTCAATGCTTGAAGAGCGGTTTG 300
 241 atgtgggtgggtgtccaaaagagcggtgacagatttcaatgcttgaagagcggtttg 300
 301 GATATTAGATACGGTGTTCGAGAAATTCATATAGTAAAGACTTGAACCTCAAGTT 360
 301 gatattagatacgggtgttcgagaaattcattatagtaaaagacttgaacctcaagtt 360
 361 GATTTTCTTACAGACGCTGGAATGCTGAAATGTTCCGAAGATCGTTTATGTCATAAA 420
 361 gattttcttaccagacgctggaatgctgaaatgttccgaagatcgtttattgcataaa 420
 421 ACATATTTAAATGGGATCATGTAAACCATCTGACTTCATGTTGTATGAGCTCTTGAT 480
 421 acatatTTAAATGGGATCATGTAAACCATCTGACTTCATGTTGTATGAGCTCTTGAT 480
 481 GTTGTATTATACATGACCAATGTGCTGCATGCGTTCCCAAAATTAAGTTTGTAA 540
 481 gttgtattatACATGACCAATGTGCTGCATGCGTTCCCAAAATTAAGTTTGTAA 540
 541 AAACGATATGAAGCCATCCCAAAATTAAGTACTGAAATCAGACAGATATATACCA 600
 541 aaacgatatGAAGCCATCCCAAAATTAAGTACTGAAATCAGACAGATATATACCA 600
 601 TGGCGTTTGAAGGGCTGGCAAGCCAGCTTTGGTGGCGACCATCTCCCAAAATTCGAT 660
 601 tggcgtttGAAGGGCTGGCAAGCCAGCTTTGGTGGCGACCATCTCCCAAAATTCGAT 660
 661 CTGGTTCGCGTGATCCCGCAATTCGCGG 691
 661 ctggttcgcggtgatcccgcaatttcgcg 691

RESULT 11
 Z58405

ID Z58405 standard; cDNA; 3618 BP.
 AC Z58405;
 DT 23-MAY-2000 (first entry)
 DE DNA encoding fumonisin esterase-amino polyol amine oxidase fusion.
 KW Amino polyol amine oxidase; APO; fumonisin esterase; mycotoxin;
 KW transgenic plant; detoxification; animal feed; silage;
 KW glutathione S transferase; alpha-amylase; signal peptide; maize;
 SS.
 OS Chimeric - Schistosoma japonicum.
 OS Chimeric - Exophiala spinifera.
 FH Key Location/Qualifiers
 FT sig_peptide 1..687
 FT /*tag= a
 FT /product= "GST - polylinker"
 FT mat_peptide 677..3615
 FT /*tag= b
 FT /product= "fumonisin-esterase-K:trAPO fusion"
 FT msc_feature 2227..2229
 FT /*tag= c
 FT /note= "extra lysine codon"
 PN WO200004159-A1.
 PD 27-JAN-2000.
 XX 08-JUL-1999; 99WO-US15454.
 XX 15-JUL-1998; 98US-0092936.
 XX 21-MAY-1999; 99US-0135391.
 XX (PION-) PIONEER HI-BRED INT INC.
 XX (CURA-) CURAGEN CORP.
 XX Duick JP, Gilliam JT, Maddox JR;
 DR WPI: 2000-182425/16.
 DR P-PSDB; Y58916.
 XX
 PT New isolated polynucleotides, polypeptides useful for detecting and
 PT degrading fumonisin or structurally related mycotoxin in processed
 PT grain or in silage
 PT
 XX
 PS Example 13; Page 115-120; 154pp; English.
 XX
 CC The present sequence is that of a polynucleotide encoding a
 CC fusion protein (see Y58916) composed of glutathione S transferase
 CC and Exophiala spinifera fumonisin esterase (ESPI) joined via a
 CC peptide linker to K:trAPO (see Y58909), a truncated, but still
 CC functional, amino polyol amine oxidase of E. spinifera. The
 CC construct was designed for expression in Escherichia coli of a hybrid
 CC protein having ESPI and APO activity. The invention provides APO
 CC polynucleotides (see Z58383-87) and polypeptides (see Y58900-05) of
 CC E. spinifera and Rhinocladiella atrovirens. The polynucleotides
 CC are used to transform plant cells normally susceptible to Fusarium
 CC or other toxin-producing fungus infection. Also provided are
 CC methods for expressing APO, optionally as a fusion protein with
 CC fumonisin esterase, in transgenic plants, prokaryotic and non-plant
 CC eukaryotic systems. Methods for detoxification of grain, grain
 CC processing, silage, food crops and in animal feed and rumen
 CC microorganisms are also disclosed.
 XX
 SQ Sequence 3618 BP; 884 A; 925 C; 957 G; 852 T; 0 other;

Query Match 62.2%; Score 682.2; DB 21; Length 3618;
 Best Local Similarity 99.6%; Pred. No. 5e-178;
 Matches 684; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

OY 1 ATGTCCTTATCTAGTATTGGAATAAAGGCGCTTGCAACCCACTGACTTCTT 60
DB 1 atgtccctactactagttatctggaataataaggcctcttgcaaccactgactctt 60
OY 61 TTGGAATATCTTGAGAAAAATATAGAGCATTTGTATGACGCGATTAAGTATAA 120
DB 61 ttggaatattcttgagaaaaatataagagcatcttgatagagcgatgaagtgataaa 120
OY 121 TGGCAACAACAAAGTTGTAATGGTTGGATTGCCAATCTTCTTATTATTATGAT 180
DB 121 tggcgaacaacaaagtgttgaaattggattggattgcccaatcttccctattatattgt 180
OY 181 GGTGATGTTAAATTAACACAGTCTATGCCATCATACGTTATATAGCTGACACACAAC 240
DB 181 ggtgatgtttaaattaacacagctctatgcatcatagctatagctgacaagcacac 240
OY 241 ATGTTGGGCTGTTGCCAAAAGAGCGTGAGAGATTTCATGCTTGAGAGCGGTTTG 300
DB 241 atgttgggctggttgccaaaagagcgctgagagatttcattgaaagcggttctg 300
OY 301 GATATTAGATACGCTGTTTCGAGAAATGCTAATAGTAAAGCTTTGAACCTCTCAAAAGTT 360
DB 301 gatattagatacggctgttccgaaattgctaaatagtaagacttgaacctccaagct 360
OY 361 GATTTCTTAGCAGCTACCTGAAAATGCTGAAAATGTTGAAAGATCGTTATGCTATAA 420
DB 361 gatcttcttagcagctacctgaatgctgaaaatgctgaaagatcgcttattgctataa 420
OY 421 ACATATTTAAATGCGATCATGTATACCCATCTGACTTATGTTATAGCGCTTGTAT 480
DB 421 acatatttaaattgctgatactgaaacccatcctgacttcaatgcttatacgccttctg 480
OY 481 GTTGTTTTATACATGACCCCAATGCTGCTGATGCGTTCGCCAAATTTAGTTTATAA 540
DB 481 gtgtttttatatactgaccccaatgctgctgattgctgttcccaaatatgattgattaaa 540
OY 541 AAAGCTATTGAAGCTATCCACAAAATGATAAGTACTTGAATAATCCAGCAATATATAGA 600
DB 541 aaagctattgaagctatcccaaaaatgataagtaacttgaataatccagcatatagca 600
OY 601 TGGCTTTGCGAGGCTGCGAACGACGTTGCTGCGGACACCTCCCAAAATGCGAT 660
DB 601 tggctttgcgaggctgcgaaacgacgcttgcgacacacctcccaaaaatgcgat 660
OY 661 CTGTTCCGCGTGATCCCGCAATTC 687
DB 661 ctggttcgcggtgatcccggaattc 687

RESULT 12
ID 260644
AC 260644:
DT 16-MAY-2000 (first entry)
XX
DE DNA encoding a fusion of aminopolylol amine oxidase/fumonisin esterase.
KW Aminopolylol amine oxidase; APO; mycotoxin degradation; antimicrobial;
  plant fungal invasion; hydrogen peroxide; Fusarium; fungi;
  fumonisin esterase; ss.
XX
OS Synthetic.
XX Exophiala spinifera.
FH Key Location/Qualifiers
FT CDS 1..3618
FT /tag= a
FT /product= "GST/aminopolylol amine oxidase/fumonisin
FT /misc-feature 1..687

```

```

FT FT /tag= b
FT FT /note= "encodes GST and a linker"
FT FT mat_peptide
FT FT 688..2190
FT FT /tag= c
FT FT /product= "fumonisin esterase"
FT FT misc-feature
FT FT 2191..2226
FT FT /tag= d
FT FT /note= "spacer sequence"
FT FT mat_peptide
FT FT 2227..3615
FT FT /tag= e
FT FT /product= "aminopolylol amine oxidase"
FT FT /note= "the protein has an amino terminal Lys
  for optimized expression"

WO200004160-A1.
PD 27-JAN-2000.
XX
XX 08-JUL-1999; 99WO-US15455.
XX
XX 15-JUL-1998; 98US-0092936.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX (CORA-) CORAGEN CORP.
XX
XX Craista OR, Duvick J, Folkerts O, Gilliam JT, Maddox JR;
XX WPI: 2000-182426/16.
XX P-PSDB: Y68852.
XX
XX New nucleic acid encoding aminopolylol amine oxidase, used, e.g. to
XX generate plants resistant to Fusarium.
XX
XX Example 13; Page 110-115; 145pp; English.
XX
XX The present sequence encodes a fusion protein of GST/Exophiala
XX spinifera aminopolylol amine oxidase (APO)/fumonisin esterase.
XX CC The APO enzyme has homology to the flavin containing amine oxidase
XX family, that oxidize primary amine to an aldehyde or ketone, releasing
XX ammonia and hydrogen peroxide. The APO enzyme degrades mycotoxins
XX CC that promote fungal invasion of plants. Destruction of mycotoxins by
XX APO generates, as a by-product, hydrogen peroxide which is itself an
XX antimicrobial and stimulates the plants own defensive systems. The
XX CC APO polynucleotides are used to generate plants (particularly maize)
XX CC that are resistant to Fusarium or other fungi that produce mycotoxins
XX and/or to degrade such mycotoxins (e.g. during ensilaging); for
XX CC recombinant production of APO polypeptides; as selection markers for
XX CC plant transformation; and to isolate related sequences from other
XX CC organisms. The APO polypeptides are used to degrade mycotoxins in
XX CC plant materials, including expression in engineered bacteria and fungi,
XX CC e.g. rumen microflora.
XX
XX Sequence 3618 BP; 884 A; 925 C; 957 G; 852 T; 0 other;

Query Match 62.2%; Score 682.2; DB 21; Length 3618;
Best Local Similarity 99.6%; Pred. No. 56-178;
Matches 684; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

DR WPI: 1995-131101/17.
DR P-PSDB: R72793.
XX
PT Nucleic acid encoding a protein or polypeptide that inhibits DNA
PT synthesis in a recipient cell - useful to inhibit cell
PT proliferation in tumour cells, treat wound or burn tissue, or as
PT antiviral or antiparasitic agent
XX
PS Disclosure: Page 129; 169pp; English.
XX
XX The senescent cell-derived inhibitor-1 (SDI-1) inhibits DNA
CC synthesis in a recipient cell and is capable of associating with a
CC cyclin (esp. cyclin D1) or cyclin-dependent kinase (esp. CDK2).
CC Fusion of SDI-1 cDNA (086773) with a gene encoding S. japonicum
CC glutathione-S-transferase (Q86774) allowed production of a fusion
CC protein from which recombinant SDI-1 was easily isolated.
XX
SQ Sequence 699 BP; 205 A; 122 C; 156 G; 216 T; 0 other;

Query Match	62.1%;	Score 681;	DB 16;	length 699;
Best Local Similarity	100.0%;	Pred. No. 5.2e-178;		
Matches 681; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0.

QY	61	TTGGAAATATCTTGAAGAAATAATGAGAGACATTTGTATGAGCGGATGCAAGCTGATATAA	120
Db	61	ttggaatattcttgaagaanaataTgaagagcatlctgataTgaagcgataTgaagTgataaa	120
QY	121	TGGCGAACAACAAAAGTTTGAATTGGGTTTGGAGTTTCCCAATCTTCTTATATATTTGAT	180
Db	121	tggcgaaacaaaagtttgaatttgggttgggttgggttcccaattcttccttattataatgat	180
QY	181	GGGAGTATTAATAATTACACACTTATATGGCATATATACGTTATATATAGCTGACACACAC	240
Db	181	gggagatgtaaatatlaaacacactlctatTgccaatactatataTgctTgacaagaacaac	240
QY	241	ATGTTGGGTGGTTTGCACAAAAGACGTCGAGAGATTTCAATGCTTGAAGGAGCGGTTTTC	300
Db	241	atgttgggtggttgcacaaaagacgctgagagattcaatgctTgaagagcggttttg	300
QY	301	GATATTACATACGCGTGTTCGAGAAATTCATATATGTAAGACTTTGAAACTCTCAAAATT	360
Db	301	gatattacatacogtgttctcgagattgcataTatagaagacttTgaactctccaagttc	360
QY	361	GATTTTCTAGACAGCTACCTGAAATGCTGTAAATTTGTGAAGATCGTTTATGTCAATAA	420
Db	361	gatttctcttagacaagctacacctgaaatgctTgaataatgtctgaagatcgcttlatgataaa	420
QY	421	ACATATTTAAATGCGTATCATATGTAAACCAATCCGACTTCATGTTGTATGAGCGCTTGAT	480
Db	421	acataTTaaatgctgataTcatgtaaacccatccgacttcaTgtglatagaagcctctgat	480
QY	481	GTGTGTTTATACATGAGACCAATGTGCTGCTGAGTGCCTTCCCAAAATTAAGTTGTTTTAA	540
Db	481	gttgttTatacatTgagaccaaatTgtgcctTgagTgcgttcccaaaattagttgttttaa	540
QY	541	AAACCTATTGAGCTATACCCACAATAATGATTAAGTACTTGAATCCAGAACTATATATGCA	600
Db	541	aaagctattgaagctatacccaacaattgataagTacttTgaaatccagaagatataTgca	600
QY	601	TGGCCTTTGCAAGGCTTGCAAGCCACGTTTGTGTGGGACACATCTCCAAAATTCGAT	660
Db	601	tggcctttTgcaaggctTgcaagccacgTttgtTgtgggacacatctccaaaatTcgat	660
QY	661	CTGGTTTCGCGCTGGATCCCG 681	
Db	661	ctggttttcgctggatcccg 681	

AA		Schistosoma japonicum.	
OS			
XX	Key	Location/Qualifiers	
PH	CDS	1..699	
FT		/tag	a
XX			
PN	W09506415-A.		
XX			
PD	09-MAR-1995.		
XX			
PF	26-AUG-1994;	94WO-US09700.	
XX			
PR	13-JUL-1994;	94US-0274535.	
PR	30-AUG-1993;	93US-0113372.	
PR	17-NOV-1993;	93US-0153564.	
PR	03-JAN-1994;	94US-0160814.	
PR	25-FEB-1994;	94US-0203535.	
PR	15-APR-1994;	94US-0229420.	
XX	30-JUN-1994;	94US-0268439.	
PA	(BAYU)	BAYLOR COLLEGE MEDICINE.	
XX			
PI	Smith JR;		
XX			

PD 28-SEP-1995.
XX
XX 20-MAR-1995; 95WO-US03191.
XX
XX 24-MAR-1994; 94US-0216894.
XX
XX (KIRC/) KIRCHHOFF L V.
PA (ORNU/) OTSU K.
XX
XX Kirchhoff LV, Otsu K;
DR WPI: 1995-344618/44.
DR P-PSDB: R84568.
XX
XX New polypeptide(s) from the TCR27 protein of Trypanosoma cruzi - as
PT immunassay reagent for specific diagnosis of Chagas disease, also
PT related nucleic acid and transformed cells
XX
XX
PS Disclosure: Page 37-39; 68pp; English.

XX T05329-T05333 encode polypeptides of the TCR27 protein of T. cruzi
CC (R84565-R84569). The proteins are all fusion products with glutathione-
CC S-transferase (GST) and may contain a linker sequence. The TCR27
CC protein comprises a 95 amino acid (aa) N-terminal region; 69 repeats
CC of a highly conserved 14 aa sequence and a 68 aa C-terminal region.
CC This sequence encodes the GST sequence, the Ag44 polypeptide contg.
CC 16 of the 69 repeat units and the amino and carboxy terminal
CC peptides of TCR27. The TCR27 polypeptides of the invention are useful
CC for the diagnosis of Chagas disease (American Trypanosomiasis), they
CC are capable of detecting anti-T. cruzi antibodies; or for blood
CC screening. The TCR27 protein has epitopes to which most T. cruzi
CC infected individuals have antibodies. The TCR27 polypeptides will not
CC react with serum from patients with leishmaniasis, schistosomiasis, or
CC autoimmune disease and are hence less likely to cause false positives
CC in diagnosis.
XX
XX

Sequence 1932 BP; 629 A; 360 C; 570 G; 373 T; 0 other;

Query Match 62.1%; Score 681; DB 16; Length 1932;

Best Local Similarity 100.0%; Pred. NO. 8.1e-178; Indels 0; Gaps 0;

Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCCCTATCTACTAGTATGTGAATAATTAGGCGCTTGCGAACCCACTGACTTCT 60
DB 1 atgtccctatactactagtatttggaataattagggccttgcgaaacccactgactctt 60
QY 61 TTGGAATATCTTGGAAGAAAATATGAAGCATTTGTATGAGCGGATGAAGTGATATA 120
DB 61 ttggaatattcttggaagaaaatatagaagcatcttgatagcgagatgaagtgataaa 120
QY 121 TGGCGAAACAAAAGTTGAATTTGGATTGGAGTTCCCAATTCCTTATATATTGAT 180
DB 121 tggcgaaacaaaagttgaatttggattggagttcccaatctccttataatattgat 180
QY 181 GGTGATGTTAAATTATACAGTCTATGGCCATCATAGCTTATATAGCTGACAAGCACAAC 240
DB 181 ggtgatgtttaaattatcacagtctatggccatcatagcttataatagctgacaagcacaa 240
QY 241 ATGTTGGGTGGTTGCCAAAAGCGTGACAGATTTCATGCTTGAAGGAGCGGTTTG 300
DB 241 atgttgggtggttgccaaaagcgtgacagatttcattgcttgaaggagcggtttg 300
QY 301 GATATTAGTAGGTTGTTGCGAATTCGATATGATAAAGACTTGAAGACTCTCAAGTT 360
DB 301 gatatttagtggttggttgctgaatttcgatatgataaagacttgaagactctcaagtt 360
QY 361 GATTTTCTTAGCAAGTACTGAAATGCTGAAGATGTTGAGAGATGTTTATGTCATATA 420
DB 361 gattttcttagcaagtactgaaatgctgaagatgttgagagatgtttattgcatataa 420
QY 421 ACATATTTAATGATGATCATGTAACCATCTGACTTCATGTTGATGAGCGCTCTGAT 480
DB 421 acatattttaatgatgatcatgttaacccatctgacttcattgttgatgagcgctctgat 480

DB 421 acatatttaaaagtgatgatalgataaccatccctgacttcatatgttgatagcgctctgat 480
QY 481 GTTGTATTATACAGCAACCAATGTGCTGATGCTTCCCAAAATAGTTGTTTAAA 540
DB 481 gttgtattatatacagcaaccaaagtgccctgatacgcttcccaaatagttgttttaa 540
QY 541 AAGGATTTGAAGCTATCCCAAAATTTGATTAAGTCTTGAATTCAGCAAGTATATAGCA 600
DB 541 aagcatattgaagctatcccaaaattgataagtaacttgaatccagcaagtatataagca 600
QY 601 TGGCTTTTGACAGGCGTGGCAAGCCACGTTTGTGTGGCGACCATCTCCCAAAATCGGAT 660
DB 601 tggccttttgacagcgctggcaagccacgtttgtgtggcgaccatctcccaaaatcgga 660
QY 661 CTGGTTCGCGGTGATGCCCG 681
DB 661 ctggttcgcggtgatgcccg 681

Search completed: March 18, 2001, 04:59:49
Job time: 2528 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2001, 03:47:41 ; Search time 56.88 Seconds

(without alignments)
3105.339 Million cell updates/sec

Title: US-09-402-488a-1

Perfect score: 1096
Sequence: 1 ATGTCCCTACTAGTGA.....ATCCACTAGTTAGACGG 1096

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 280836 seqs, 80580151 residues

Total number of hits satisfying chosen parameters: 561672

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/PCRB_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	681	62.1	699	4	PCT-US94-09700-5
2	681	62.1	1065	2	US-08-216-894-3
3	681	62.1	1695	2	US-08-216-894-1
4	681	62.1	1932	2	US-08-216-894-7
5	680	62.0	1194	4	PCT-US94-09700-10
6	678.4	61.9	1281	3	US-08-737-248-3
7	678.2	61.9	1740	2	US-08-864-224-10
8	667.4	60.9	897	1	US-08-102-757-8
9	663.4	60.5	1419	2	US-08-216-894-9
10	663.4	60.5	1920	3	US-08-974-180-19
11	662.8	60.5	2095	2	US-08-305-764C-55
12	661	60.3	924	2	US-08-216-894-5
13	660.4	60.3	1951	3	US-08-895-707-1
14	654	59.7	654	1	US-08-470-837-25
15	649	59.2	897	1	US-08-102-757-10
16	151.8	11.9	1458	1	US-07-854-596B-42
17	151.8	11.9	1467	1	US-07-854-596B-46
18	151.6	11.8	201	1	US-07-854-596B-1
19	151.6	11.8	223	1	US-07-854-596B-7
20	151.6	11.8	420	1	US-07-854-596B-8
21	144.2	11.2	1082	1	US-08-715-252-1
22	144.2	11.2	1102	2	US-08-453-051-3
23	144.2	11.2	1130	1	US-08-406-948A-5
24	140.6	12.8	217	5	5164304-3
25	140.2	12.8	227	5	5422249-13
26	139.4	12.7	563	1	US-08-044-506B-22
27	138.6	12.6	206	1	US-08-044-506B-21
28	138	12.6	195	1	US-08-406-948A-7

29	137.8	12.6	279	1	US-08-186-222-3	Sequence 3, Appl
30	137.6	12.6	195	1	US-08-044-506B-27	Sequence 27, Appl
31	137	12.5	202	1	US-08-044-506B-20	Sequence 20, Appl
32	136.4	12.4	195	1	US-08-044-506B-28	Sequence 28, Appl
33	131.2	12.0	212	5	5180668-10	Patent No. 5180668
34	131.2	12.0	227	2	US-07-982-064-8	Sequence 8, Appl
35	131.2	12.0	8491	2	US-08-757-439-1	Sequence 1, Appl
36	129.8	11.8	2115	2	US-08-767-026-3	Sequence 3, Appl
37	129	11.8	1240	1	US-08-240-372-2	Sequence 2, Appl
38	127.6	11.6	218	5	5422249-4	Patent No. 5422249
39	127	11.6	2732	5	5217891-14	Patent No. 5217891
40	127	11.6	2733	2	US-08-846-021A-6	Sequence 6, Appl
41	126.8	11.6	235	1	US-08-262-384A-7	Sequence 7, Appl
42	126.8	11.6	235	1	US-08-262-384A-8	Sequence 8, Appl
43	126.8	11.6	726	1	US-08-262-384A-9	Sequence 9, Appl
44	126.8	11.6	726	1	US-08-262-384A-10	Sequence 10, Appl
45	125.8	11.5	198	1	US-07-910-528-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
PCT-US94-09700-5
Sequence 5, Application PC/TUS9409700
GENERAL INFORMATION:
APPLICANT: BAYLOR COLLEGE OF MEDICINE
APPLICANT: SMITH, JAMES R.
TITLE OF INVENTION: SENESCENT CELL DERIVED INHIBITORS OF
TITLE OF INVENTION: DNA SYNTHESIS
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESSES:
ADDRESSEE: HOWREY & SIMON
STREET: 1299 PENNSYLVANIA AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09700
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/808,523
FILING DATE: 16-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/970,462
FILING DATE: 02-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/113,372
FILING DATE: 30-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/153,564
FILING DATE: 17-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/203,535
FILING DATE: 25-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/229,420
FILING DATE: 15-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/274,535
FILING DATE: 13-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: AUERBACH, JEFFREY I.
REGISTRATION NUMBER: 32,680
REFERENCE/DOCKET NUMBER: 225-102-CIP7-PCT
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 383-7451
 TELEFAX: (202) 383-6610
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 699 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Schistosoma japonicum
 IMMEDIATE SOURCE:
 CLONE: GST
 PCT-US94-09700-5

Query Match 62.1%; Score 681; DB 4; Length 699;
 Best Local Similarity 100.0%; Pred. No. 9.3e-185;
 Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGCCCTTACTAGTGGTTATTTGAAATTAAGGCTTGTGCAACCCACTGACTTCTT 60
 DB 1 ATGCCCTTACTAGTGGTTATTTGAAATTAAGGCTTGTGCAACCCACTGACTTCTT 60
 OY 61 TTGGAATCTTTGAAGAAATTAAGAGCATTTGTATGAGCGGATGAAGGTATATA 120
 DB 61 TTGGAATCTTTGAAGAAATTAAGAGCATTTGTATGAGCGGATGAAGGTATATA 120
 OY 121 TGGCAAAACAAATTTGAAATTTGAGTTTCCCAATCTTCTATATATATATAT 180
 DB 121 TGGCAAAACAAATTTGAAATTTGAGTTTCCCAATCTTCTATATATATATATAT 180
 OY 181 GGTGATGTTAAATTAACACAGTCTATGCGCATCATATATATATATATATATAT 240
 DB 181 GGTGATGTTAAATTAACACAGTCTATGCGCATCATATATATATATATATATATAT 240
 OY 241 ATGTGGGGGTGGTTCGCAAAAGAGGTGAGAGTTTCAATCTTGAAGGAGGCTTTTG 300
 DB 241 ATGTGGGGGTGGTTCGCAAAAGAGGTGAGAGTTTCAATCTTGAAGGAGGCTTTTG 300
 OY 301 GATATTAGATACGTTGTTGAGAAATTCATATAGTAAAGACTTTGAAACTCTCAAGTT 360
 DB 301 GATATTAGATACGTTGTTGAGAAATTCATATAGTAAAGACTTTGAAACTCTCAAGTT 360
 OY 361 GATTTTCTTAGACAGCTTACCTGAAATGCTGAAATGTTGGAAGATCGTTATGATATA 420
 DB 361 GATTTTCTTAGACAGCTTACCTGAAATGCTGAAATGTTGGAAGATCGTTATGATATA 420
 OY 421 ACATATTTAAATGATGATCATGTAACCCATCTGACTTCTGATGATGACCTCTTGAT 480
 DB 421 ACATATTTAAATGATGATCATGTAACCCATCTGACTTCTGATGATGACCTCTTGAT 480
 OY 481 GTTGTGTTTATACATGAGCCAAATGTCCTGATGCGTTGCCAAATTTAGTTTGTATAA 540
 DB 481 GTTGTGTTTATACATGAGCCAAATGTCCTGATGCGTTGCCAAATTTAGTTTGTATAA 540
 OY 541 AAACGATTTGAAGCTATCCCAAAATTTGTAAGTACTTAAATCCAGCAAGATATATAGA 600
 DB 541 AAACGATTTGAAGCTATCCCAAAATTTGTAAGTACTTAAATCCAGCAAGATATATAGA 600
 OY 601 TGGCCCTTTCAGAGGCTGGCAAGCCAGCTTTGGTGGTGGGAGCATTCCTCAAAATCGAT 660
 DB 601 TGGCCCTTTCAGAGGCTGGCAAGCCAGCTTTGGTGGTGGGAGCATTCCTCAAAATCGAT 660
 OY 661 CTGGTTCCCGCTGGATCCCG 681
 DB 661 CTGGTTCCCGCTGGATCCCG 681

RESULT 2
 US-08-216-894-3

Sequence 3, Application US/08216894
 Patent No. 5876734
 GENERAL INFORMATION:
 APPLICANT: Kitchhoff, Louis V.
 APPLICANT: Otsu, Keiko
 TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION
 TITLE OF INVENTION: WITH TRYPAOSOMA CRUZI
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Foley & Lardner
 STREET: 3000 K Street, N.W., Suite 500
 CITY: Washington, D.C.
 COUNTRY: USA
 ZIP: 20007-5109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/216,894
 FILING DATE: 24-MAR-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 85326/102/DRL0
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)672-5300
 TELEFAX: (202)672-5399
 TELEX: 904136
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1065 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1062
 US-08-216-894-3

Query Match 62.1%; Score 681; DB 2; Length 1065;
 Best Local Similarity 100.0%; Pred. No. 1.1e-184;
 Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGCCCTTACTAGTGGTTATTTGAAATTAAGGCTTGTGCAACCCACTGACTTCTT 60
 DB 1 ATGCCCTTACTAGTGGTTATTTGAAATTAAGGCTTGTGCAACCCACTGACTTCTT 60
 OY 61 TTGGAATCTTTGAAGAAATTAAGAGCATTTGTATGAGCGGATGAAGGTATATA 120
 DB 61 TTGGAATCTTTGAAGAAATTAAGAGCATTTGTATGAGCGGATGAAGGTATATA 120
 OY 121 TGGCAAAACAAATTTGAAATTTGAGTTTCCCAATCTTCTATATATATATATAT 180
 DB 121 TGGCAAAACAAATTTGAAATTTGAGTTTCCCAATCTTCTATATATATATATATAT 180
 OY 181 GGTGATGTTAAATTAACACAGTCTATGCGCATCATATATATATATATATATATAT 240
 DB 181 GGTGATGTTAAATTAACACAGTCTATGCGCATCATATATATATATATATATATATAT 240
 OY 241 ATGTGGGGGTGGTTCGCAAAAGAGGTGAGAGTTTCAATCTTGAAGGAGGCTTTTG 300
 DB 241 ATGTGGGGGTGGTTCGCAAAAGAGGTGAGAGTTTCAATCTTGAAGGAGGCTTTTG 300
 OY 301 GATATTAGATACGTTGTTGAGAAATTCATATAGTAAAGACTTTGAAACTCTCAAGTT 360
 DB 301 GATATTAGATACGTTGTTGAGAAATTCATATAGTAAAGACTTTGAAACTCTCAAGTT 360
 OY 361 GATTTTCTTAGACAGCTTACCTGAAATGCTGAAATGTTGGAAGATCGTTATGATATAA 420
 DB 361 GATTTTCTTAGACAGCTTACCTGAAATGCTGAAATGTTGGAAGATCGTTATGATATAA 420

Db 361 GATTTCTTAGACAGCTACCTGAATGCTGAAATGTTGGAAGTCTTATGTCTATAA 420
Qy 421 ACATATTTAAATGATGATCATGTAACCAATCCCTGACTGATGTTATGACGCTTGAT 480
Db 421 ACATATTTAAATGATGATCATGTAACCAATCCCTGACTGATGTTATGACGCTTGAT 480
Qy 481 GTTGTATTAACATGAGCAACCAATGTCCTGATGCTGCTCCAAATTAATGTTGTTTAA 540
Db 481 GTTGTATTAACATGAGCAACCAATGTCCTGATGCTGCTCCAAATTAATGTTGTTTAA 540
Qy 541 AAACGATTTGAAGCTATGCCACAATTAATGATTAATTAATCCGCAAGTATATAGCA 600
Db 541 AAACGATTTGAAGCTATGCCACAATTAATGATTAATTAATCCGCAAGTATATAGCA 600
Qy 601 TGCGCTTTTGACAGGCTGCGCAAGCAGCTTTGTGTGTGCGCAGCAGTCTCCAAATGCGAT 660
Db 601 TGCGCTTTTGACAGGCTGCGCAAGCAGCTTTGTGTGTGCGCAGCAGTCTCCAAATGCGAT 660
Qy 661 CTGGTTCCGCGTGATCCCG 681
Db 661 CTGGTTCCGCGTGATCCCG 681

RESULT 3

US-08-216-894-1
; Sequence 1, Application US/08216894
; Patent No. 5876734
; GENERAL INFORMATION:
; APPLICANT: Kirchoff, Louis V.
; APPLICANT: Otsu, Keiko
; TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION
; TITLE OF INVENTION: WITH TRYPAOSOMA CRUZI
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/216,894
; FILING DATE: 24-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 85326/102/DRLO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1695 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1692
; US-08-216-894-1

Query Match 62.1%; Score 681; DB 2; Length 1695;
Best Local Similarity 100.0%; Pred. No. 1.4e-184;
Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTCCCTATACCTAGCTATTGGAATAATTAGAGGCGCTTGTCAGCAACCACTCGACTTCTT 60

Db 1 ATGTCCCTATACCTAGCTATTGGAATAATTAGAGGCGCTTGTCAGCAACCACTCGACTTCTT 60
Qy 61 TTGGATATCTTGAAGAAAATATGSAAGCATTTCTATAGAGCGCATGAAGTGATATA 120
Db 61 TTGGATATCTTGAAGAAAATATGSAAGCATTTCTATAGAGCGCATGAAGTGATATA 120
Qy 121 TGCGCAACAAAAGTTGAATTTGGGTTTGAGTTTCCCAATCTTCTATTAATTAATGAT 180
Db 121 TGCGCAACAAAAGTTGAATTTGGGTTTGAGTTTCCCAATCTTCTATTAATTAATGAT 180
Qy 181 GTGATGTTAAATTAACACAGTCTATGCGCATATACGTTATATAGCTGACAGCAAC 240
Db 181 GTGATGTTAAATTAACACAGTCTATGCGCATATACGTTATATAGCTGACAGCAAC 240
Qy 241 ATGTGGGTGGTTGTCCAAAGAGCGTGCAGAGATTTCAATGCTTAAGAGAGCGTTTG 300
Db 241 ATGTGGGTGGTTGTCCAAAGAGCGTGCAGAGATTTCAATGCTTAAGAGAGCGTTTG 300
Qy 301 GATATTAGATACGCTGTTTCGAGAAATTCATATAGTAAGACTTTGAACTCTCAAGTT 360
Db 301 GATATTAGATACGCTGTTTCGAGAAATTCATATAGTAAGACTTTGAACTCTCAAGTT 360
Qy 361 GATTTCTTAGACAGCTACCTGAATGCTGAAAATGTTGAGAGTCTTTATGTCTATAA 420
Db 361 GATTTCTTAGACAGCTACCTGAATGCTGAAAATGTTGAGAGTCTTTATGTCTATAA 420
Qy 421 ACATATTTAAATGATGATATGTAACCAATCCCTGACTGATGTTATGATACGCTTGAT 480
Db 421 ACATATTTAAATGATGATATGTAACCAATCCCTGACTGATGTTATGATACGCTTGAT 480
Qy 481 GTTGTATTAACATGAGCAACCAATGTCCTGATGCTGCTCCAAATTAATGTTGTTTAA 540
Db 481 GTTGTATTAACATGAGCAACCAATGTCCTGATGCTGCTCCAAATTAATGTTGTTTAA 540
Qy 541 AAACGATTTGAAGCTATGCCACAATTAATGATTAATTAATCCGCAAGTATATAGCA 600
Db 541 AAACGATTTGAAGCTATGCCACAATTAATGATTAATTAATCCGCAAGTATATAGCA 600
Qy 601 TGCGCTTTTGACAGGCTGCGCAAGCAGCTTTGTGTGTGCGCAGCAGTCTCCAAATGCGAT 660
Db 601 TGCGCTTTTGACAGGCTGCGCAAGCAGCTTTGTGTGTGCGCAGCAGTCTCCAAATGCGAT 660
Qy 661 CTGGTTCCGCGTGATCCCG 681
Db 661 CTGGTTCCGCGTGATCCCG 681

RESULT 4

US-08-216-894-7
; Sequence 7, Application US/08216894
; Patent No. 5876734
; GENERAL INFORMATION:
; APPLICANT: Kirchoff, Louis V.
; APPLICANT: Otsu, Keiko
; TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION
; TITLE OF INVENTION: WITH TRYPAOSOMA CRUZI
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/216,894
; FILING DATE: 24-MAR-1994
; ATTORNEY/AGENT INFORMATION:

```

1 NAME: BENT, Stephen A.
2 REGISTRATION NUMBER: 29, 768
3 REFERENCE/DOCKET NUMBER: 85326/102/DRUC
4 TELECOMMUNICATION INFORMATION:
5 TELEPHONE: (202)672-5300
6 TELEFAX: (202)672-5399
7
8 TELEFAX: 904136
9
10 INFORMATION FOR SEQ ID NO: 7:
11
12 SEQUENCE CHARACTERISTICS:
13
14 LENGTH: 1932 base pairs
15 TYPE: nucleic acid
16 STRANDEDNESS: double
17 TOPOLOGY: linear
18
19 MOLECULE TYPE: DNA (genomic)
20
21 FEATURE:
22
23 NAME/KEY: CDS
24 LOCATION: 1..1929
25
26 OS-08-216-894-7

```

Query Match	Score	DB 2;	Length	1932;
62.18;	681;			

QY	1	ATGTCCTTAACATCTGGTTATATGGAATAATTAAGGACCTGTGTCAACCCACTGACTCTT	60
Db	1	ATGTCCCTATACATCTGGTTATATGGAATAATTAAGGACCTGTGTCAACCCACTGACTCTT	60
QY	61	TTGGAATATCTTGAAGAAAAATATGAAGACATTTGTATAGCCGGATGAAGGTGATAA	120
Db	61	TTGGAATATCTTGAAGAAAAATATGAAGACCTTTGTATAGACCGGATGAAGGTGATAA	120
QY	121	TGGGCAACCAAAACCTTGAATTGGGTTTGGAGTTTCCCAATCTTCCTATATATATGAT	180
Db	121	TGGGCAACCAAAACCTTGAATTGGGTTTGGAGTTTCCCAATCTTCCTATATATATGAT	180
QY	181	GGTGAATGTTAAATTTAACAACAGTCTATGGCCATCATACGTTATATAGCTGACAAGACAAC	240
Db	181	GGTGAATGTTAAATTTAACAACAGTCTATGGCCATCATACGTTATATAGCTGACAAGACAAC	240
QY	241	ATGTTGGGATGGTGTGCCAAAAGAGGTCGAGAGATTTCAATGCTTGAAGAGCGGTTTG	300
Db	241	ATGTTGGGATGGTGTGCCAAAAGAGGTCGAGAGATTTCAATGCTTGAAGAGCGGTTTG	300
QY	301	GATATTAGATACGGGTGTTGAGAAATTCGATATAGTAAGAAGCTTTGAAACTGTCAAAATT	360
Db	301	GATATTAGATACGGGTGTTGAGAAATTCGATATAGTAAGAAGCTTTGAAACTGTCAAAATT	360
QY	361	GATTTTCTTAGCAACCTACCTGAATGCTGAAAATGTTGGAAGATCGTTATATGCTATAA	420
Db	361	GATTTTCTTAGCAACCTACCTGAATGCTGAAAATGTTGGAAGATCGTTATATGCTATAA	420
QY	421	ACATATTAAATGGAGATCATGTAAACCCATCCCTGACTCATGTGTGATGACGCTCTGAT	480
Db	421	ACATATTAAATGGAGATCATGTAAACCCATCCCTGACTCATGTGTGATGACGCTCTGAT	480
QY	481	GTTGTTTATATACATGAGACCAATGTGCTGAGATGCGTATCCCAAAATTAAGTTGTTTAA	540
Db	481	GTTGTTTATATACATGAGACCAATGTGCTGAGATGCGTATCCCAAAATTAAGTTGTTTAA	540
QY	541	AAACGATTTGAAGCTATCCCACAAAATGTATAGTACTTGAATCCGACAGTATATAGCA	600
Db	541	AAACGATTTGAAGCTATCCCACAAAATGTATAGTACTTGAATCCGACAGTATATAGCA	600
QY	601	TGGCCTTTGACAGGCTGGGCAAGCACAGTTGTGTGTGAGGACCATCTCCAAAATGGGAT	660
Db	601	TGGCCTTTGACAGGCTGGGCAAGCACAGTTGTGTGTGAGGACCATCTCCAAAATGGGAT	660
QY	661	CTGGTTCCGGCTGGATCCCCG 681	
Db	661	CTGGTTCCGGCTGGATCCCCG 681	

RESULT 5
PCT-US94-09700-10
Sequence 10. Application PC/TUS9409700
GENERAL INFORMATION:
APPLICANT: BAYLOR COLLEGE OF MEDICINE
APPLICANT: SMITH, JAMES R.
TITLE OF INVENTION: SENSECENT CELL DERIVED INHIBITORS OF
TITLE OF INVENTION: DNA SYNTHESIS
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOWREY & SIMON
STREET: 1299 PENNSYLVANIA AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM type: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09700
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/808,523
FILING DATE: 16-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/970,462
FILING DATE: 02-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/113,372
FILING DATE: 30-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/153,564
FILING DATE: 17-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/203,535
FILING DATE: 25-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/229,420
FILING DATE: 15-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/274,535
FILING DATE: 13-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: AUERBACH, JEFFREY I.
REGISTRATION NUMBER: 32,680
REFERENCE/DOCKET NUMBER: 225-102-CIP7-PC7
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 383-7451
TELEFAX: (202) 383-6610
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1194 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: GST-SD1-1 gene fusion
PCT-US94-09700-10

Query Match	62.0%	Score 680	DB 4	Length 1194
Best Local Similarity	100.0%	Pred. NO. 2.3e-184		
Matches 680	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY	1	ATGTCCCTATACGTGGTTATTGGAAAAATTAAGGCGCTTGACACCAACGACTTCCTT	60

Db 1 ATGTCCCTATACGTATTTGGAATTTAGGGCTGTGCAACCCACTGCTTCT 60
Qy 61 TTGGATATCTTGAAGAAAATATGAGAGCATTTGTATGAGCGGATGAGTGAATA 120
Db 61 TTGGATATCTTGAAGAAAATATGAGAGCATTTGTATGAGCGGATGAGTGAATA 120
Qy 121 TGGCGAACAACAAAGTTGAATTTGGGTGGAGTTCCCAATCTCTTATATTTAT 180
Db 121 TGGCGAACAACAAAGTTGAATTTGGGTGGAGTTCCCAATCTCTTATATTTAT 180
Qy 181 GGTGATGTTAAATTAACAGCTCTATGCGCATCATCTTATATGCTGACAGCAAC 240
Db 181 GGTGATGTTAAATTAACAGCTCTATGCGCATCATCTTATATGCTGACAGCAAC 240
Qy 241 ATGTGGGTGGTGTCCAAAGAGCGTGCAGAGATTTCAATGCTTGAAGAGCGTTTG 300
Db 241 ATGTGGGTGGTGTCCAAAGAGCGTGCAGAGATTTCAATGCTTGAAGAGCGTTTG 300
Qy 301 GATATTAGTACGGTGTTCGAGAAATGATATGTAAGACTTTGAACCTCAAGTT 360
Db 301 GATATTAGTACGGTGTTCGAGAAATGATATGTAAGACTTTGAACCTCAAGTT 360
Qy 361 GATTTCTTAGCAAGCTACCTGAATGCTGAAAATGTTGAGAGATGCTTATGCTTAA 420
Db 361 GATTTCTTAGCAAGCTACCTGAATGCTGAAAATGTTGAGAGATGCTTATGCTTAA 420
Qy 421 ACATATTTAAATGTTGATGATGATTAACCATCTGACTTCATGTTATGAGCGTCTT 480
Db 421 ACATATTTAAATGTTGATGATGATTAACCATCTGACTTCATGTTATGAGCGTCTT 480
Qy 481 GTTGTTTTATACATGAGCAATGTGCTGATGCTTCCCAAAATTTAGTTGTTTAA 540
Db 481 GTTGTTTTATACATGAGCAATGTGCTGATGCTTCCCAAAATTTAGTTGTTTAA 540
Qy 541 AAACGATTGAAGCTATCCACAAATTTGATTAAGTACTTGAATCCAGCAATATAGCA 600
Db 541 AAACGATTGAAGCTATCCACAAATTTGATTAAGTACTTGAATCCAGCAATATAGCA 600
Qy 601 TGGCCCTTTCAGAGGCTGCGCAAGCAGCTTTGTTGTTGGGACCATCTCCAAATCGAT 660
Db 601 TGGCCCTTTCAGAGGCTGCGCAAGCAGCTTTGTTGTTGGGACCATCTCCAAATCGAT 660
Qy 661 CTGGTTCGCGTGGATCCCC 680
Db 661 CTGGTTCGCGTGGATCCCC 680

RESULT 6
US-08-737-248-3
Sequence 3, Application US/08737248
Patent No. 6114305
GENERAL INFORMATION:
APPLICANT: Guemene, Daniel
APPLICANT: zadworny, David
APPLICANT: Karalzas, Costas
TITLE OF INVENTION: USE OF PROLACTIN FOR PREVENTING AND/OR
NUMBER OF SEQUENCES: 23
TREATING BIRD BROODINESS
CORRESPONDENCE ADDRESS:
ADDRESSEE: WEISER & ASSOCIATES
STREET: 230 South Fifteenth Street, Suite 500
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentia Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/737,248
FILING DATE: 28-APR-1997

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR95/00576
FILING DATE: 03-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 94/05550
FILING DATE: 05-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weiser, Gerard J.
REGISTRATION NUMBER: 19,763
REFERENCE/DOCKET NUMBER: 989,6411P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-875-8393
TELEFAX: 215-875-8394
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1281 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1278
US-08-737-248-3

Query Match 61.9%; Score 678.4; DB 3; Length 1281;
Best Local Similarity 99.9%; Pred. No. 6.8e-184;
Matches 679; Conservative 0; Mismatches 1; Indels 0; Gaps 0:

Qy 1 ATGTCCCTATACGTATTTGGAATTTAGGGCTGTGCAACCCACTGCTTCT 60
Db 1 ATGTCCCTATACGTATTTGGAATTTAGGGCTGTGCAACCCACTGCTTCT 60
Qy 61 TTGGATATCTTGAAGAAAATATGAGAGCATTTGTATGAGCGGATGAGTGAATA 120
Db 61 TTGGATATCTTGAAGAAAATATGAGAGCATTTGTATGAGCGGATGAGTGAATA 120
Qy 121 TGGCGAACAACAAAGTTGAATTTGGGTGGAGTTCCCAATCTCTTATATTTAT 180
Db 121 TGGCGAACAACAAAGTTGAATTTGGGTGGAGTTCCCAATCTCTTATATTTAT 180
Qy 181 GGTGATGTTAAATTAACAGCTCTATGCGCATCATCTTATATGCTGACAGCAAC 240
Db 181 GGTGATGTTAAATTAACAGCTCTATGCGCATCATCTTATATGCTGACAGCAAC 240
Qy 241 ATGTGGGTGGTGTCCAAAGAGCGTGCAGAGATTTCAATGCTTGAAGAGCGTTTG 300
Db 241 ATGTGGGTGGTGTCCAAAGAGCGTGCAGAGATTTCAATGCTTGAAGAGCGTTTG 300
Qy 301 GATATTAGTACGGTGTTCGAGAAATGATATGTAAGACTTTGAACCTCAAGTT 360
Db 301 GATATTAGTACGGTGTTCGAGAAATGATATGTAAGACTTTGAACCTCAAGTT 360
Qy 361 GATTTCTTAGCAAGCTACCTGAATGCTGAAAATGTTGAGAGATGCTTATGCTTAA 420
Db 361 GATTTCTTAGCAAGCTACCTGAATGCTGAAAATGTTGAGAGATGCTTATGCTTAA 420
Qy 421 ACATATTTAAATGTTGATGATGATTAACCATCTGACTTCATGTTATGAGCGTCTT 480
Db 421 ACATATTTAAATGTTGATGATGATTAACCATCTGACTTCATGTTATGAGCGTCTT 480
Qy 481 GTTGTTTTATACATGAGCAATGTGCTGATGCTTCCCAAAATTTAGTTGTTTAA 540
Db 481 GTTGTTTTATACATGAGCAATGTGCTGATGCTTCCCAAAATTTAGTTGTTTAA 540
Qy 541 AAACGATTGAAGCTATCCACAAATTTGATTAAGTACTTGAATCCAGCAATATAGCA 600
Db 541 AAACGATTGAAGCTATCCACAAATTTGATTAAGTACTTGAATCCAGCAATATAGCA 600
Qy 601 TGGCCCTTTCAGAGGCTGCGCAAGCAGCTTTGTTGTTGGGACCATCTCCAAATCGAT 660
Db 601 TGGCCCTTTCAGAGGCTGCGCAAGCAGCTTTGTTGTTGGGACCATCTCCAAATCGAT 660

Db 601 TGCGCTTTCAGGGCTGCGCAAGCCAGCTTTGGTGTGCGAGCAATCTCCAAATGGAT 660
QY 661 CTGGTTCGGCTGATCCCC 680
|||||
Db 661 CTGGTTCGGCTGATCCCC 680

RESULT 7

US-08-864-224-10

Sequence 10, Application US/08864224

Patent No. 5851808

GENERAL INFORMATION:

APPLICANT: Eileedge, Stephen J.

APPLICANT: Liu, Qinhua

TITLE OF INVENTION: Rapid Subcloning Using Site-Specific

TITLE OF INVENTION: Recombination

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

ADDRESSEE: Medlen & Carroll, LLP

STREET: 220 Montgomery Street, Suite 2200

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/864,224

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Ingolia, Diane E.

REGISTRATION NUMBER: 40,027

REFERENCE/DOCKET NUMBER: BCM-02681

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 1740 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid

DESCRIPTION: /desc = "DNA"

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1737

US-08-864-224-10

Query Match 61.9%; Score 678.2; DB 2; Length 1740;
Best Local Similarity 98.1%; Pred. No. 9e-184;
Matches 686; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 ATGCCCTATCTAGTATTTGGAATTAAGGCCCTTGTCACCCAGCTGATCTTCT 60
|||||
Db 1 ATGCCCTATCTAGTATTTGGAATTAAGGCCCTTGTCACCCAGCTGATCTTCT 60
61 TTGGAATATCTTGAAGAAATATGACGATTTGTATGACCGGATGAAGTATATA 120
|||||
Db 61 TTGGAATATCTTGAAGAAATATGACGATTTGTATGACCGGATGAAGTATATA 120
121 TGGCAAAACAAAATTTGAATTTGGGTTTGAAGTTTCCCAATCTTCTATTATTGAT 180
|||||
Db 121 TGGCAAAACAAAATTTGAATTTGGGTTTGAAGTTTCCCAATCTTCTATTATTGAT 180
121 TGGCAAAACAAAATTTGAATTTGGGTTTGAAGTTTCCCAATCTTCTATTATTGAT 180
|||||
QY 181 GGTATGTTAAATTAACACAGCTATGCGCATCATGCTTATATAGTGACAAAC 240
|||||
Db 181 GGTATGTTAAATTAACACAGCTATGCGCATCATGCTTATATAGTGACAAAC 240

QY 241 ATGTGGGTGGTGTGTCAAAAGAGCGTCAGAGATTTCATGCTTGAAGGAGCGGTTTG 300
|||||
Db 241 ATGTGGGTGGTGTGTCAAAAGAGCGTCAGAGATTTCATGCTTGAAGGAGCGGTTTG 300
301 GATATTAGATACGGTGTTCGAGAAATTCATATAGTAAAGACTTTGAACCTTCAAAGTT 360
|||||
Db 301 GATATTAGATACGGTGTTCGAGAAATTCATATAGTAAAGACTTTGAACCTTCAAAGTT 360
301 GATATTAGATACGGTGTTCGAGAAATTCATATAGTAAAGACTTTGAACCTTCAAAGTT 360
|||||
QY 361 GATTTCTTACCAAGCTACCTGAATTCGTAAGAAATGTCGAAGATCGTTTATGATATA 420
|||||
Db 361 GATTTCTTACCAAGCTACCTGAATTCGTAAGAAATGTCGAAGATCGTTTATGATATA 420
361 GATTTCTTACCAAGCTACCTGAATTCGTAAGAAATGTCGAAGATCGTTTATGATATA 420
|||||
QY 421 ACATATTTAAATGATGATATGTAACCCATCCCTGATGATGTTTATGACGCTTGTAT 480
|||||
Db 421 ACATATTTAAATGATGATATGTAACCCATCCCTGATGATGTTTATGACGCTTGTAT 480
421 ACATATTTAAATGATGATATGTAACCCATCCCTGATGATGTTTATGACGCTTGTAT 480
|||||
QY 481 GTTGTATTATACATGAGACCAATGCTGCTGATGCTGCTCCCAAAATTTGTTTATA 540
|||||
Db 481 GTTGTATTATACATGAGACCAATGCTGCTGATGCTGCTCCCAAAATTTGTTTATA 540
541 AACGATTTGAAGCTATCCCAATTTGATAGTACTTGAATCCAGCATATATAGCA 600
|||||
QY 541 AACGATTTGAAGCTATCCCAATTTGATAGTACTTGAATCCAGCATATATAGCA 600
541 AACGATTTGAAGCTATCCCAATTTGATAGTACTTGAATCCAGCATATATAGCA 600
|||||
QY 601 TGCCCTTTCAGAGGCTGCGCAAGCCAGCTTTGGTGTGCGCAGCATCTCCAAATCGGAT 660
|||||
Db 601 TGCCCTTTCAGAGGCTGCGCAAGCCAGCTTTGGTGTGCGCAGCATCTCCAAATCGGAT 660
601 TGCCCTTTCAGAGGCTGCGCAAGCCAGCTTTGGTGTGCGCAGCATCTCCAAATCGGAT 660
|||||
QY 661 CTGGTTCGGCTGATCCCCGGAATTCGCGGATCGACTG 699
|||||
Db 661 CTGGTTCGGCTGATCCCCGGAATTCGCGGATCGACTG 699

RESULT 8

US-08-102-757-8

Sequence 8, Application US/08102757

Patent No. 5480864

GENERAL INFORMATION:

APPLICANT: Tajima, Masahiro

APPLICANT: Ohnuma, Mamami

APPLICANT: Lerner, Ethan A.

TITLE OF INVENTION: Modified Maxadilan Protein, Its

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &

ADDRESSEE: Dunner

STREET: 1300 I Street, N.W., Suite 700

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/102,757

FILING DATE: 01-JAN-1993

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Meyers, Kenneth J.

REGISTRATION NUMBER: 25,146

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-408-4000

TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 897 base pairs

TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-102-757-8

Query Match 60.9%; Score 667.4; DB 1; Length 897;
Best Local Similarity 99.1%; Pred. No. 7.8e-181;
Matches 671; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 ATGTCCTATCTAGTATGGAATTAAGGCGCTTGCCACCCAGCTGACTCTT 60
DB 1 ATGTCCTATCTAGTATGGAATTAAGGCGCTTGCCACCCAGCTGACTCTT 60
OY 61 TTGGAATATCTTGAAGAAATATGAGAGCATTTGTATGAGCGGATGAAGTATAA 120
DB 61 TTGGAATATCTTGAAGAAATATGAGAGCATTTGTATGAGCGGATGAAGTATAA 120
OY 121 TGGCCAAACAAAGTTGAAATGGGTTTGAGCTTCCCAATCTTCTATTATTCAT 180
DB 121 TGGCCAAACAAAGTTGAAATGGGTTTGAGCTTCCCAATCTTCTATTATTCAT 180
OY 181 GGTGATGTTAAATTAACAGCTATGCGCATCATACGTATATAGCTGACAGCAAC 240
DB 181 GGTGATGTTAAATTAACAGCTATGCGCATCATACGTATATAGCTGACAGCAAC 240
OY 241 ATGTTGGGTGGTTGCCAAAAGAGCGTGAGAGATTTCAATCTTGAAGAGCGGTTTG 300
DB 241 ATGTTGGGTGGTTGCCAAAAGAGCGTGAGAGATTTCAATCTTGAAGAGCGGTTTG 300
OY 301 GATATTAGATACGTTGTTGCAAAATTCGATTAAGACCTTTGAAACCTCTCAAGTT 360
DB 301 GATATTAGATACGTTGTTGCAAAATTCGATTAAGACCTTTGAAACCTCTCAAGTT 360
OY 361 GATTTCTTACAGACTACCTGAATGCTGAAATGTTGGAAGATGTTTATGCTATAA 420
DB 361 GATTTCTTACAGACTACCTGAATGCTGAAATGTTGGAAGATGTTTATGCTATAA 420
OY 421 ACATATTTAAATGATGATCATGTATTAACCCATCTGACTTATGTTATGAGCGCTTGAT 480
DB 421 ACATATTTAAATGATGATCATGTATTAACCCATCTGACTTATGTTATGAGCGCTTGAT 480
OY 481 GTTGTATTAATACATGAGCCCAATGTCCTGATGCGTTTCCCAAAATTAAGTTTAA 540
DB 481 GTTGTATTAATACATGAGCCCAATGTCCTGATGCGTTTCCCAAAATTAAGTTTAA 540
OY 541 AAACGATTTGAAGCTATCCCAAAATGATAGTACTTGAATCCAGCAAGTATATACCA 600
DB 541 AAACGATTTGAAGCTATCCCAAAATGATAGTACTTGAATCCAGCAAGTATATACCA 600
OY 601 TGGCCTTTGACAGGCTGGCAAGCCAGCTTGGTGGTGGCGACATCTCCAAATGCGAT 660
DB 601 TGGCCTTTGACAGGCTGGCAAGCCAGCTTGGTGGTGGCGACATCTCCAAATGCGAT 660
OY 661 CTGTTCCGCGGTGATC 677
DB 661 CTGTTCCGCGGTGATC 677

RESULT 9
US-08-216-894-9
Sequence 9, Application US/08216894
Patent No. 5876734

GENERAL INFORMATION:
APPLICANT: Kirchhoff, Louis V.
APPLICANT: Otsu, Keiko
TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION
TITLE OF INVENTION: WITH TRYPAOSOMA CRUZI
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.

COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/216, 894
FILING DATE: 24-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 85326/102/DRLO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1419 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1416
US-08-216-894-9

Query Match 60.5%; Score 663.4; DB 2; Length 1419;
Best Local Similarity 99.8%; Pred. No. 1.3e-179;
Matches 664; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ATGTCCTATCTAGTATGGAATTAAGGCGCTTGCCACCCAGCTGACTCTT 60
DB 1 ATGTCCTATCTAGTATGGAATTAAGGCGCTTGCCACCCAGCTGACTCTT 60
OY 61 TTGGAATATCTTGAAGAAATATGAGAGCATTTGTATGAGCGGATGAAGTATAA 120
DB 61 TTGGAATATCTTGAAGAAATATGAGAGCATTTGTATGAGCGGATGAAGTATAA 120
OY 121 TGGCCAAACAAAGTTGAAATGGGTTTGAGCTTCCCAATCTTCTATTATTCAT 180
DB 121 TGGCCAAACAAAGTTGAAATGGGTTTGAGCTTCCCAATCTTCTATTATTCAT 180
OY 181 GGTGATGTTAAATTAACAGCTATGCGCATCATACGTATATAGCTGACAGCAAC 240
DB 181 GGTGATGTTAAATTAACAGCTATGCGCATCATACGTATATAGCTGACAGCAAC 240
OY 241 ATGTTGGGTGGTTGCCAAAAGAGCGTGAGAGATTTCAATGCTTGAAGAGCGGTTTG 300
DB 241 ATGTTGGGTGGTTGCCAAAAGAGCGTGAGAGATTTCAATGCTTGAAGAGCGGTTTG 300
OY 301 GATATTAGATACGTTTTCGAAATGTCATATAGTAAGACTTTGAAACTCAAGTT 360
DB 301 GATATTAGATACGTTTTCGAAATGTCATATAGTAAGACTTTGAAACTCAAGTT 360
OY 361 GATTTCTTACAGACTACCTGAATGCTGAAATGCTGAAAGCTTTATGCTATAA 420
DB 361 GATTTCTTACAGACTACCTGAATGCTGAAATGCTGAAAGCTTTATGCTATAA 420
OY 421 ACATATTTAAATGATGATCATGTATTAACCCATCTGACTTATGTTATGAGCGCTTGAT 480
DB 421 ACATATTTAAATGATGATCATGTATTAACCCATCTGACTTATGTTATGAGCGCTTGAT 480
OY 481 GTTGTATTAATACATGAGCCCAATGTCCTGATGCGTTCCCAAAATTAAGTTTAA 540
DB 481 GTTGTATTAATACATGAGCCCAATGTCCTGATGCGTTCCCAAAATTAAGTTTAA 540
OY 541 AAACGATTTGAAGCTATCCCAAAATGATAGTACTTGAATCCAGCAAGTATATACCA 600
DB 541 AAACGATTTGAAGCTATCCCAAAATGATAGTACTTGAATCCAGCAAGTATATACCA 600

Db 541 AACGCTATGAACTATCCACAAATGATAGTACTGAAATCCAGCAAGTATATAGCA 600

QY 601 TGGCCTTTGACGGGCTGGCAAGCCAGTTGGTGTGGCCACATCCCAAAATGGAT 660
|||||

Db 601 TGGCCTTTGACGGGCTGGCAAGCCAGTTGGTGTGGCCACATCCCAAAATGGAT 660
|||||

QY 661 CTGGT 665
|||

Db 661 CTGAT 665

RESULT 10
US-08-974-180-19

Sequence 19, Application US/08974180

Patent No. 6025194

GENERAL INFORMATION:

APPLICANT: Funk, Walter

TITLE OF INVENTION: Methods for Modulating and Identifying

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genon Corporation

STREET: 230 Constitution Drive

CITY: Menlo Park

STATE: California

COUNTRY: USA

ZIP: 94025

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/974,180

FILING DATE: 19-NOV-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Kaster, Kevin R.

REGISTRATION NUMBER: 32,704

REFERENCE/DOCKET NUMBER: 206

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 473-7779

TELEFAX: (650) 473-8654

INFORMATION FOR SEQ. ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 1920 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

FEATURE:

NAME/KEY: CDS

LOCATION: 258..1868

OTHER INFORMATION: /product="recombinant GS-GC6

OTHER INFORMATION: fusion protein"

US-08-974-180-19

Query Match 60.5%; Score 663.4; DB 3; Length 1920;
Best Local Similarity 99.8%; Pred. No. 1.6e-179;
Matches 664; Conservative 0; Mismatches 1; Indels 0; Caps 0;

QY 1 ATGTCCCTATAGTATGTAATAAGGCTGTGCAACCCACTGACTTCTT 60
|||||

Db 258 ATGTCCCTATAGTATGTAATAAGGCTGTGCAACCCACTGACTTCTT 317
|||||

QY 61 TTGGAATATCTGAAGAAATATGAGAGCATTTGATAGCCGATGAAGTGATAA 120
|||||

Db 318 TTGGAATATCTGAAGAAATATGAGAGCATTTGATAGCCGATGAAGTGATAA 377
|||||

QY 121 TGGGAAACAAAAGTTGATGGTTGGAGTTGCCAATCTTCTTATATATGAT 180
|||||

Db 378 TGGGAAACAAAAGTTGATGGTTGGAGTTGCCAATCTTCTTATATATGAT 437
|||||

QY 181 GGTGATGTTAAATTAACACAGTCTATGGCCATCATAGCTTATATAGTCGACAGACAAC 240
|||||

Db 438 GGTGATGTTAAATTAACACAGTCTATGGCCATCATAGCTTATATAGTCGACAGACAAC 497
|||||

QY 241 ATGTGGGTGGTGTTCACAAAGAGCGTCAGAGATTTCAATGCTGGAAGAGCGGTTTG 300
|||||

Db 438 ATGTGGGTGGTGTTCACAAAGAGCGTCAGAGATTTCAATGCTGGAAGAGCGGTTTG 557
|||||

QY 301 GATATTACATGAGGTGTTTCGAGAAATTCATATAGCTTAAGACTTGAACCTCAAGATT 360
|||||

Db 558 GATATTACATGAGGTGTTTCGAGAAATTCATATAGCTTAAGACTTGAACCTCAAGATT 617
|||||

QY 361 GATTTCTTACAGACCTGCAATGCTGAATAAGTTGCAAGATCGTTTGTGCTATAA 420
|||||

Db 618 GATTTCTTACAGACCTGCAATGCTGAATAAGTTGCAAGATCGTTTGTGCTATAA 677
|||||

QY 421 ACATATTTAAATGATGATCATGTAACCCATCTGATGTTGATGACGCTTGAT 480
|||||

Db 678 ACATATTTAAATGATGATCATGTAACCCATCTGATGTTGATGACGCTTGAT 737
|||||

QY 481 GTTGTATATCATGAGACCATGTCCTGATGCTGCCAATAATTAAGTTGTTTAA 540
|||||

Db 738 GTTGTATATCATGAGACCATGTCCTGATGCTGCCAATAATTAAGTTGTTTAA 797
|||||

QY 541 AACGCTATGAACTATCCACAAATGATAGTACTGAAATCCAGCAAGTATATAGCA 600
|||||

Db 798 AACGCTATGAACTATCCACAAATGATAGTACTGAAATCCAGCAAGTATATAGCA 857
|||||

QY 601 TGGCCTTTGACGGGCTGGCAAGCCAGTTGGTGTGGCCACATCCCAAAATGGAT 660
|||||

Db 858 TGGCCTTTGACGGGCTGGCAAGCCAGTTGGTGTGGCCACATCCCAAAATGGAT 917
|||||

QY 661 CTGGT 665
|||

Db 918 CTGAT 922

RESULT 11
US-08-305-764C-55

Sequence 55, Application US/08305764C

Patent No. 5856090

GENERAL INFORMATION:

APPLICANT: Epstein, David M.

TITLE OF INVENTION: DNA METHYLASE LINKING REACTION

NUMBER OF SEQUENCES: 71

CORRESPONDENCE ADDRESS:

ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE

STREET: 10550 No. 5856090th Torrey Pines Road

CITY: La Jolla

STATE: California

COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/305,764C

FILING DATE: 09-SEP-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Fitting, Thomas

REGISTRATION NUMBER: 34,163

REFERENCE/DOCKET NUMBER: TSRI 440.0

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 784-2937

TELEFAX: (619) 784-9399

INFORMATION FOR SEQ. ID NO: 55:

SEQUENCE CHARACTERISTICS:

LENGTH: 2095 base pairs

TYPE: nucleic acid

STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 9..1952
US-08-305-764C-55

Query Match 60.5%; Score 662.8; DB 2; Length 2095;
Best Local Similarity 99.0%; Pred. No. 2.4e-179;
Matches 667; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

1 ATGTCCCTATCTAGTATTGGAATAAAGGCGCTGTGCAACCCCTGACCTCTT 60
9 ATGTCCCTATCTAGTATTGGAATAAAGGCGCTGTGCAACCCCTGACCTCTT 68
61 TTGGAATATCTGAAAGAAAATATGAAGAGCATTTGTATAGCGCGATGAGTGATAA 120
69 TTGGAATATCTGAAAGAAAATATGAAGAGCATTTGTATAGCGCGATGAGTGATAA 128
121 TGGCGAAGAAAAGTTGATTTGGTTGGAGTTTCCCAATCTCTTATTATTGAT 180
129 TGGCGAAGAAAAGTTGATTTGGTTGGAGTTTCCCAATCTCTTATTATTGAT 188
181 GGTGATGTTAAATTAACAGCTATAGGCATCATAGCTTATATAGCTGACAGCAAC 240
189 GGTGATGTTAAATTAACAGCTATAGGCATCATAGCTTATATAGCTGACAGCAAC 248
241 ATGTGGGTGTTGTCACAAAGAGCGTGAGATTTCAATGCTTGAAGAGCGTTTGG 300
249 ATGTGGGTGTTGTCACAAAGAGCGTGAGATTTCAATGCTTGAAGAGCGTTTGG 308
301 GATATTAGTACGCTGTTTGCAGATTTGATTAAGACTTTGAAACTCTCAAGTT 360
309 GATATTAGTACGCTGTTTGCAGATTTGATTAAGACTTTGAAACTCTCAAGTT 368
361 GATTTTCTAGCAAGCTACCTGAATGCTGAATGCTTGAAGATGTTTATGCTATAA 420
369 GATTTTCTAGCAAGCTACCTGAATGCTGAATGCTTGAAGATGTTTATGCTATAA 428
421 ACATATTTAAATGCTGATCATGTAAACCATCTGACTTCAATGTTATGAGCTTTGAT 480
429 ACATATTTAAATGCTGATCATGTAAACCATCTGACTTCAATGTTATGAGCTTTGAT 488
481 GTTGTTTTATACATGAGCAACCATGTCCTGATGCTTCCCAAAATTAAGTTTATAA 540
489 GTTGTTTTATACATGAGCAACCATGTCCTGATGCTTCCCAAAATTAAGTTTATAA 548
541 AAACGATTTGAAGCTATCCACAAATGATAGTACTTGAATCCGCAAGTTATAGCA 600
549 AAACGATTTGAAGCTATCCACAAATGATAGTACTTGAATCCGCAAGTTATAGCA 608
601 TGGCCTTTGACAGGCTGCGCAAGCAGTGTGTTGGTGGGACCATCTCCAAATGAGAT 660
609 TGGCCTTTGACAGGCTGCGCAAGCAGTGTGTTGGTGGGACCATCTCCAAATGAGAT 668
661 CTGGTTCCGCTGG 674
669 ATCGAAGTCTGG 682

RESULT 12
US-08-216-894-5
Sequence 5, Application US/08216894
Patent No. 5876734
GENERAL INFORMATION:
APPLICANT: Kirchoff, Louis V.
APPLICANT: Otsu, Keiko
TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION WITH TRYPAOSOMA CRUZI

NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/216,894
FILING DATE: 24-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 85326/102/DRLO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 924 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..921
US-08-216-894-5

Query Match 60.3%; Score 661; DB 2; Length 924;
Best Local Similarity 100.0%; Pred. No. 5.3e-179;
Matches 661; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGTCCCTATCTAGTATTGGAATAAAGGCGCTGTGCAACCCCTGACCTCTT 60
1 ATGTCCCTATCTAGTATTGGAATAAAGGCGCTGTGCAACCCCTGACCTCTT 60
61 TTGGAATATCTGAAAGAAAATATGAAGAGCATTTGTATAGCGCGATGAGTGATAA 120
61 TTGGAATATCTGAAAGAAAATATGAAGAGCATTTGTATAGCGCGATGAGTGATAA 120
121 TGGCGAAGAAAAGTTGATTTGGTTGGAGTTTCCCAATCTCTTATTATTGAT 180
121 TGGCGAAGAAAAGTTGATTTGGTTGGAGTTTCCCAATCTCTTATTATTGAT 180
181 GGTGATGTTAAATTAACAGCTATAGGCATCATAGCTTATATAGCTGACAGCAAC 240
181 GGTGATGTTAAATTAACAGCTATAGGCATCATAGCTTATATAGCTGACAGCAAC 240
241 ATGTGGGTGTTGTCACAAAGAGCGTGAGATTTCAATGCTTGAAGAGCGTTTGG 300
241 ATGTGGGTGTTGTCACAAAGAGCGTGAGATTTCAATGCTTGAAGAGCGTTTGG 300
301 GATATTAGTACGCTGTTTGCAGATTTGATTAAGACTTTGAAACTCTCAAGTT 360
301 GATATTAGTACGCTGTTTGCAGATTTGATTAAGACTTTGAAACTCTCAAGTT 360
361 GATTTTCTAGCAAGCTACCTGAATGCTGAATGCTTGAAGATGTTTATGCTATAA 420
361 GATTTTCTAGCAAGCTACCTGAATGCTGAATGCTTGAAGATGTTTATGCTATAA 420
421 ACATATTTAAATGCTGATCATGTAAACCATCTGACTTCAATGTTATGAGCTTTGAT 480
421 ACATATTTAAATGCTGATCATGTAAACCATCTGACTTCAATGTTATGAGCTTTGAT 480
481 GTTGTTTTATACATGAGCAACCATGTCCTGATGCTTCCCAAAATTAAGTTTATAA 540

|||||
Db 481 GTGTGTTATATACATGAGCCCAATGCTGATGCTGCCAAATATGTTTAA 540
|||
Qy 541 AACGATATGAGCTATATCCCAATATGATAGTATGAAATCCAGCATATATACA 600
|||
Db 541 AACGATATGAGCTATATCCCAATATGATAGTATGAAATCCAGCATATATACA 600
|||
Qy 601 TGGCCTTTGACAGGGCTGCGAAGCCAGTGTGTGTGCGAGCATCTCCAAATCGGAT 660
|||
Db 601 TGGCCTTTGACAGGGCTGCGAAGCCAGTGTGTGTGCGAGCATCTCCAAATCGGAT 660
|||
Qy 661 C 661
|||
Db 661 C 661
|||
RESULT 13
US-08-895-707-1
; Sequence 1, Application US/08895707
; Patent No. 6077700
GENERAL INFORMATION:
APPLICANT: (Pharmacia & Upjohn, Co.)
APPLICANT: alternatively, for U.S. Filing:
APPLICANT: Hollingsworth, Robert A.
APPLICANT: Sharma, Satish K.
APPLICANT: Rank, Kenneth B.
APPLICANT: Evans, David B.
TITLE OF INVENTION: Special Constructs and Complexes of
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESS: Pharmacia & Upjohn Company
STREET: 301 Henrietta Street
CITY: Kalamazoo
STATE: MI
COUNTRY: USA
ZIP: 49001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Rel. #1.0, Ver. #1.25/WordPerfect 5.2+
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/895,707
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Woollon, Thomas A.
REGISTRATION NUMBER: 35,004
REFERENCE/DOCKET NUMBER: 6054
TELECOMMUNICATION INFORMATION:
TELEPHONE: 616-833-7914
TELEFAX: 616-833-8897
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1951 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-895-707-1

Query Match 60.3%; Score 660.4; DB 3; Length 1951;
Best Local Similarity 99.8%; Pred. No. 1.1e-178;
Matches 661; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGTCCCATATCTAGTATTGGAATAATAGGGCTTGTGCAACCCACTGACTTCT 60
|||
Db 1 ATGTCCCATATCTAGTATTGGAATAATAGGGCTTGTGCAACCCACTGACTTCT 60

Qy 61 TTGAATATCTTGAAGAAAATATGAGAGCATTTGTATGAGCCGATGATATAA 120
|||
Db 61 TTGAATATCTTGAAGAAAATATGAGAGCATTTGTATGAGCCGATGATATAA 120
|||
Qy 121 TGGCGAACAACAAGTTGAAATGGGTTTGAGATTTCCCAATCTTCTATATATGAT 180
|||
Db 121 TGGCGAACAACAAGTTGAAATGGGTTTGAGATTTCCCAATCTTCTATATATGAT 180
|||
Qy 181 GGTATGTTAATATACACATCTATGCGCATCATAGCTTTATATAGCGACAACACAC 240
|||
Db 181 GGTATGTTAATATACACATCTATGCGCATCATAGCTTTATATAGCGACAACACAC 240
|||
Qy 241 ATGTTGGGTGCTGTGCAAGAGCGTCAGAGATTTCAATGCTTGAAGAGCGGTTTG 300
|||
Db 241 ATGTTGGGTGCTGTGCAAGAGCGTCAGAGATTTCAATGCTTGAAGAGCGGTTTG 300
|||
Qy 301 GATATTTAGATACGCTGTGTCAGAAATGTCATATATAGTAAGACTTGAACCTCAAGCT 360
|||
Db 301 GATATTTAGATACGCTGTGTCAGAAATGTCATATATAGTAAGACTTGAACCTCAAGCT 360
|||
Qy 361 GATTTCTTACAGCTACCTGAAATGCTGAAATGCTGAAATGCTTATGATATAA 420
|||
Db 361 GATTTCTTACAGCTACCTGAAATGCTGAAATGCTGAAATGCTTATGATATAA 420
|||
Qy 421 ACATATTTAAATGCTGATCATATGTAACCCATCTGACTTCATGTTGATAGAGCTCTGAT 480
|||
Db 421 ACATATTTAAATGCTGATCATATGTAACCCATCTGACTTCATGTTGATAGAGCTCTGAT 480
|||
Qy 481 GTGTTTATATACATGAGCCCAATGCTGCTGATGCTGCTCCCAATTTGTTTAA 540
|||
Db 481 GTGTTTATATACATGAGCCCAATGCTGCTGATGCTGCTCCCAATTTGTTTAA 540
|||
Qy 541 AACGATATGAGCTATATCCCAATATGATAGTATGAAATCCAGCAAGATATATACA 600
|||
Db 541 AACGATATGAGCTATATCCCAATATGATAGTATGAAATCCAGCAAGATATATACA 600
|||
Qy 601 TGGCCTTTGACAGGGCTGCGAAGCCAGTGTGTGTGCGAGCATCTCCAAATCGGAT 660
|||
Db 601 TGGCCTTTGACAGGGCTGCGAAGCCAGTGTGTGTGCGAGCATCTCCAAATCGGAT 660
|||
Qy 661 CT 662
|||
Db 661 CT 662
|||

RESULT 14
US-08-470-837-25
; Sequence 25, Application US/08470837
; Patent No. 5800811
GENERAL INFORMATION:
APPLICANT: Nimol, Marcel E.
APPLICANT: Hall, Frederick L.
APPLICANT: Tuan, Tai-Lan
APPLICANT: Wu, Lingtao
APPLICANT: Cheung, David T.
TITLE OF INVENTION: Transforming Growth Factor B Fusion
TITLE OF INVENTION: and
TITLE OF INVENTION: Their Use in Wound Healing
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant & Gould
STREET: 1150 Santa Monica Boulevard, Suite 400
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90025-3395
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/470,837
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharp, Janice A.
REGISTRATION NUMBER: 34,051
REFERENCE/DOCKET NUMBER: 30630-1US01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 310-445-1140
TELEFAX: 310-445-9031
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 654 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..654
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1
US-08-470-837-25

Query Match 59.7%; Score 654; DB 1; Length 654;
Best Local Similarity 100.0%; Pred. No. 4.4e-177;
Matches 654; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGTCCCTATCTAGCTATTGGAATAATGAGGCTTGTCACACCTGACTCTT 60
1 ATGTCCCTATCTAGCTATTGGAATAATGAGGCTTGTCACACCTGACTCTT 60
61 TTGGAATATCTTGAAGAAAATATGAAGAGCATTTGATGAGCGGATGAGGTAA 120
61 TTGGAATATCTTGAAGAAAATATGAAGAGCATTTGATGAGCGGATGAGGTAA 120
61 TTGGAATATCTTGAAGAAAATATGAAGAGCATTTGATGAGCGGATGAGGTAA 120
121 TGGCGAACAACAAAGTTGATGAGTTGGAGTTTCCCATCTTCTATATATGAT 180
121 TGGCGAACAACAAAGTTGATGAGTTGGAGTTTCCCATCTTCTATATATGAT 180
121 TGGCGAACAACAAAGTTGATGAGTTGGAGTTTCCCATCTTCTATATATGAT 180
181 GGTGATGTTAAATTAACACAGCTCATAGCCATCATAGCTTATATAGCTGACAAC 240
181 GGTGATGTTAAATTAACACAGCTCATAGCCATCATAGCTTATATAGCTGACAAC 240
241 ATGTGGGTGTTGCCAAGAGCGTGAGATTTCAATGCTTAAAGAGCGTTTG 300
241 ATGTGGGTGTTGCCAAGAGCGTGAGATTTCAATGCTTAAAGAGCGTTTG 300
301 GATATTAGTAGCGGTGTTGAGAAATGATATAGTAAGACTTTGAACTCTCAAGTT 360
301 GATATTAGTAGCGGTGTTGAGAAATGATATAGTAAGACTTTGAACTCTCAAGTT 360
361 GATTTTCTAGACAGCTACCTGAATGCTGAAAATGTTGAGAGATGTTATATGATAAA 420
361 GATTTTCTAGACAGCTACCTGAATGCTGAAAATGTTGAGAGATGTTATATGATAAA 420
421 ACATATTAAATGATGATGATGAACCATCCGACTTATGTTATAGACGCTTGAT 480
421 ACATATTAAATGATGATGATGAACCATCCGACTTATGTTATAGACGCTTGAT 480
481 GTTGTATTATACATGAGCAACCATGTGCTGATGCTGCCAATAATGATGTTTAA 540
481 GTTGTATTATACATGAGCAACCATGTGCTGATGCTGCCAATAATGATGTTTAA 540
541 AAACGATTGAAGCTATCCCAAAATTTGATAGTACTTGAATCCGACAGTATATAGCA 600
541 AAACGATTGAAGCTATCCCAAAATTTGATAGTACTTGAATCCGACAGTATATAGCA 600
601 TGGCCTTTGACAGGCTGCGAAGCAGCTTTGGTGGGAGACCATCTCCAAA 654
601 TGGCCTTTGACAGGCTGCGAAGCAGCTTTGGTGGGAGACCATCTCCAAA 654

RESULT 15
US-08-102-757-10
Sequence 10, Application US/08102757
Patent No. 5480864
GENERAL INFORMATION:
APPLICANT: Tajima, Masahiro
APPLICANT: Ohnuma, Mamami
APPLICANT: Lerner, Ethan A.
TITLE OF INVENTION: Modified Maxadilan Protein, Its
PREPARATION AND USE, AND DNA ENCODING THE PROTEIN
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegun, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/102,757
FILING DATE: 01-JAN-1993
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 05136.0003-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4400
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 897 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-102-757-10

Query Match 59.2%; Score 649; DB 1; Length 897;
Best Local Similarity 98.5%; Pred. No. 1.4e-175;
Matches 655; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

1 ATGTCCCTATCTAGCTATTGGAATAATGAGGCTTGTCACACCTGACTCTT 60
1 ATGTCCCTATCTAGCTATTGGAATAATGAGGCTTGTCACACCTGACTCTT 60
61 TTGGAATATCTTGAAGAAAATATGAAGAGCATTTGATGAGCGGATGAGGTAA 120
61 TTGGAATATCTTGAAGAAAATATGAAGAGCATTTGATGAGCGGATGAGGTAA 120
61 TTGGAATATCTTGAAGAAAATATGAAGAGCATTTGATGAGCGGATGAGGTAA 120
121 TGGCGAACAACAAAGTTGATGAGTTGGAGTTTCCCATCTTCTATATATGAT 180
121 TGGCGAACAACAAAGTTGATGAGTTGGAGTTTCCCATCTTCTATATATGAT 180
181 GGTGATGTTAAATTAACACAGCTCATAGCCATCATAGCTTATATAGCTGACAAC 240
181 GGTGATGTTAAATTAACACAGCTCATAGCCATCATAGCTTATATAGCTGACAAC 240
241 ATGTGGGTGTTGCCAAGAGCGTGAGATTTCAATGCTTAAAGAGCGGTTTG 300
241 ATGTGGGTGTTGCCAAGAGCGTGAGATTTCAATGCTTAAAGAGCGGTTTG 300
301 GATATTAGTAGCGGTGTTGAGAAATGATATAGTAAGACTTTGAACTCTCAAGTT 360
301 GATATTAGTAGCGGTGTTGAGAAATGATATAGTAAGACTTTGAACTCTCAAGTT 360

```
Db 301 GATATTAGATACGGTGTGTTGAGAAATGCATATAGTAAGACTTTGAAACTCTCAAGTT 360
QY 361 GATTTCTTAGCAAGCTACCTGAATGCTGAATAATGTTGAGACATGCTTATGTCATTA 420
    |||
Db 361 GATTTCTTAGCAAGCTACCTGAATGCTGAATAATGTTGAGACATGCTTATGTCATTA 420
QY 421 ACATATTAAATGCTGATCATGTAAACCATCTGACTTCATGTTGATGACGCTTTGAT 480
    |||
Db 421 ACATATTAAATGCTGATCATGTAAACCATCTGACTTCATGTTGATGACGCTTTGAT 480
QY 481 GTGTTTATACATGAGACCAATGTCCTGGATGCGTTCGCAAAATTAGTTGTTTAA 540
    |||
Db 481 GTGTTTATACATGAGACCAATGTCCTGGATGCGTTCGCAAAATTAGTTGTTTAA 540
QY 541 AAACGTATTGAACCTATCCACAAATTGATTAAGTACTTGAATCCAGCAAGTATATAGCA 600
    |||
Db 541 AAACGTATTGAACCTATCCACAAATTGATTAAGTACTTGAATCCAGCAAGTATATAGCA 600
QY 601 TGCCCTTTGCAAGGCTGGCAAGCCACGTTTGTGTGGGACCATCTCCAAATGCGAT 660
    |||
Db 601 TGCCCTTTGCAAGGCTGGCAAGCCACGTTTGTGTGGGACCATCTCCAAATGCGAT 660
QY 661 CTGGT 665
    |||
Db 661 CTGAT 665
```

Search completed: March 18, 2001, 04:58:09
Job time: 4228 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2001, 03:21:31 ; Search time 885.87 Seconds
(without alignments)
8669.677 Million cell updates/sec

Title: US-09-402-488A-1
Perfect score: 1096
Sequence: 1 ATGTCCTTACTAGCTTA.....ATCCACTAGTTCTAGACCGC 1096

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7991742 segs, 3503743858 residues

Total number of hits satisfying chosen parameters: 15983484

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: em_estba:*
42: em_estfm:*
43: em_esthum1:*

44: em_esthum2:*
45: em_esthum3:*
46: em_esthum4:*
47: em_esthum5:*
48: em_esthum6:*
49: em_esthum7:*
50: em_esthum8:*
51: em_esthum9:*
52: em_esthum10:*
53: em_esthum11:*
54: em_esthum12:*
55: em_esthum13:*
56: em_esthum14:*
57: em_esthum15:*
58: em_esthum16:*
59: em_esthum17:*
60: em_esthum18:*
61: em_esthum19:*
62: em_esthum20:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estov1:*
68: em_estov2:*
69: em_estp11:*
70: em_estp12:*
71: em_estp13:*
72: em_estp14:*
73: em_estp15:*
74: em_estro1:*
75: em_estro2:*
76: em_estro3:*
77: em_estro4:*
78: em_estro5:*
79: em_estro6:*
80: em_estro7:*
81: em_estro8:*
82: em_estro9:*
83: em_estro10:*
84: em_estro11:*
85: em_estro12:*
86: em_estro13:*
87: gb_est41:*
88: gb_est42:*
89: gb_est43:*
90: gb_est44:*
91: gb_est45:*
92: gb_est46:*
93: gb_est47:*
94: gb_est48:*
95: gb_est49:*
96: gb_est50:*
97: gb_est51:*
98: gb_est52:*
99: gb_est53:*
100: gb_est54:*
101: gb_est55:*
102: gb_est56:*
103: gb_est57:*
104: gb_est58:*
105: gb_est59:*
106: gb_est60:*
107: gb_est70:*
108: gb_est71:*
109: gb_est72:*
110: gb_est73:*
111: gb_est74:*
112: em_esthum21:*
113: em_esthum22:*
114: em_esthum23:*
115: em_estom1:*
116: em_estom2:*

117: em_estp16:*
 118: em_estp17:*
 119: em_estp18:*
 120: em_estp19:*
 121: em_estp20:*
 122: em_estp21:*
 123: em_estp22:*
 124: em_estp23:*
 125: em_estp24:*
 126: em_estp25:*
 127: em_estp26:*
 128: em_estp27:*
 129: em_estp28:*
 130: em_estp29:*
 131: em_estp30:*
 132: em_estp31:*
 133: em_estp32:*
 134: em_estp33:*
 135: em_estp34:*
 136: em_estp35:*
 137: em_estp36:*
 138: em_estp37:*
 139: em_estp38:*
 140: em_estp39:*
 141: em_estp40:*
 142: em_estp41:*
 143: em_estp42:*
 144: em_estp43:*
 145: em_estp44:*
 146: em_estp45:*
 147: em_estp46:*
 148: em_estp47:*
 149: em_estp48:*
 150: em_estp49:*
 151: em_estp50:*
 152: em_estp51:*
 153: em_estp52:*
 154: em_estp53:*
 155: em_estp54:*
 156: em_estp55:*
 157: em_estp56:*
 158: em_estp57:*
 159: em_estp58:*
 160: em_estp59:*
 161: em_estp60:*
 162: em_estp61:*
 163: em_estp62:*
 164: em_estp63:*
 165: em_estp64:*
 166: em_estp65:*
 167: em_estp66:*
 168: em_estp67:*
 169: em_estp68:*
 170: em_estp69:*
 171: em_estp70:*
 172: em_estp71:*
 173: em_estp72:*
 174: em_estp73:*
 175: em_estp74:*
 176: em_estp75:*
 177: em_estp76:*
 178: em_estp77:*
 179: em_estp78:*
 180: em_estp79:*
 181: em_estp80:*
 182: em_estp81:*
 183: em_estp82:*
 184: em_estp83:*
 185: em_estp84:*
 186: em_estp85:*
 187: em_estp86:*
 188: em_estp87:*
 189: em_estp88:*

190: gb_gss25:*
 191: gb_gss26:*
 192: gb_gss27:*
 193: gb_gss28:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	636	58.0	668	87	AM191462	AM191462 JAA000493
2	607.4	55.4	662	27	AI932072	AI932072 JAA000214
3	600.2	54.8	601	87	AM186543	AM186543 JAA000443
4	510	46.5	551	87	AM231225	AM231225 JAVH0009.
5	445.2	40.6	669	38	AM017133	AM017133 EST272555
6	420.6	38.4	636	38	AM017155	AM017155 EST272577
7	363	33.1	475	27	AI976995	AI976995 EST271589
8	329.4	30.1	421	89	AM329883	AM329883 JAT10240.
9	229.6	20.9	368	15	AI068236	AI068236 EST209927
10	223.2	20.4	438	14	AA999498	AA999498 MCE3400.
11	219.8	20.1	238	15	AI083054	AI083054 JAA00053.
12	219.6	20.0	828	20	AI446859	AI446859 EPM5034 E
13	127.8	11.7	550	93	AM654220	AM654220 103527 MA
14	127	11.6	624	88	AM239463	AM239463 x04008.Y
15	119	10.9	544	37	AV590629	AV590629 AV590629
16	111.6	10.2	601	37	AV594046	AV594046 AV594046
17	110	10.0	552	135	BE752939	BE752939 205854 MA
18	109.6	10.0	550	38	AV667937	AV667937 AV667937
19	108.6	9.9	582	109	BE506028	BE506028 d018408.Y
20	107	9.8	825	12	AA792550	AA792550 v869103.T
21	105.4	9.6	671	25	AI797367	AI797367 w87c09.X
22	102.8	9.4	868	23	AI663649	AI663649 u199h08.Y
23	102.6	9.4	676	135	BE727497	BE727497 601560236
24	102.4	9.3	964	146	W34862	W34862 mc62e02.r1
25	102.2	9.3	794	16	AI128211	AI128211 g647h02.x
26	101.2	9.2	523	93	AM652611	AM652611 100786 MA
27	101	9.2	544	111	BE682854	BE682854 181175 MA
28	100.8	9.2	651	14	AA968299	AA968299 uc71910.Y
29	100.6	9.2	574	105	BE232478	BE232478 137501 MA
30	99.6	9.1	806	23	AI663617	AI663617 u199e07.Y
31	98.8	9.0	703	2	AA137683	AA137683 mc31g06.r
32	98.4	9.0	702	2	AA110288	AA110288 mp32d07.r
33	98.4	9.0	705	107	BE383177	BE383177 601298726
34	98.4	9.0	731	107	BE386146	BE386146 601277269
35	98.4	9.0	805	135	BE783081	BE783081 601470668
36	98.2	9.0	726	3	AA199001	AA199001 mv39d05.r
37	97.6	8.9	603	8	AA498061	AA498061 v175b01.r
38	97.6	8.9	665	10	AA701663	AA701663 z143905.S
39	97.6	8.9	884	17	AI226687	AI226687 u114e02.Y
40	97.2	8.9	787	26	AI870867	AI870867 w177c10.X
41	96.8	8.8	576	28	AL134603	AL134603 DKFZP547E
42	96.4	8.8	597	107	BE392895	BE392895 601308718
43	95.2	8.7	500	108	BE486040	BE486040 173388 BA
44	95	8.7	525	95	AM823858	AM823858 uf60d11.Y
45	94.8	8.6	680	137	BE912561	BE912561 601666193

ALIGNMENTS

RESULT 1
 AM191462
 LOCUS JAA000493.SH3 Adult Sjc 7/94 Schistosoma japonicum cDNA 5', mRNA
 DEFINITION
 sequence.
 ACCESSION AM191462
 VERSION AM191462.1 GI:6467090
 KEYWORDS EST.
 SOURCE Schistosoma japonicum.

20-MAR-2000

ORGANISM Schistosoma japonicum
 Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea
 1 Strigeidae; Schistosomatidae; Schistosomatidae; Schistosoma.
 AUTHORS Hu, W., Brindley, P.J., and Feng, Z.
 TITLE Expressed sequence tags from adults of Schistosoma japonicum (Anhui strain) (Hu, Brindley, Feng)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Brindley, P.J.
 Molecular Parasitology Unit
 Queensland Institute of Medical Research
 300 Herston Road, Queensland 4029, Australia
 Tel: 61 7 3362 0413
 Fax: 61 7 3362 0104
 Email: paul.beglmr.edu.au
 PCR Primers
 FORWARD: M13 Forward
 BACKWARD: M13 Reverse
 Seq primer: T3 Reverse
 High quality sequence stop: 547.
 Location/Qualifiers

FEATURES
 source 1..668
 /organism="Schistosoma japonicum"
 /strain="Chinese (Anhui) strain"
 /db_xref="taxon:6182"
 /clone_lib="Adult SJC 7/94"
 /sex="Male and female"
 /tissue_type="Whole body"
 /dev_stage="Adult worms"
 /lab_host="Mouse and rabbit"
 /note="Vector: Lambda ZAP-11 XR; Site 1: EcoR I; Site 2: XhoI I; Several hundred adult Schistosoma japonicum (Anhui, P. R. China, strain), of mixed sex, were perfused from the mesenteries of experimentally infected mice and rabbits at the Queensland Institute of Medical Research, Brisbane, Australia (QIMR), and stored for several months in liquid nitrogen. Subsequently, mRNA was isolated at the QIMR from lysates of these worms by oligo dt chromatography, using a kit from Pharmacia. The mRNA was then shipped to Clontech, Palo Alto, CA, USA, who constructed a cDNA library. First strand synthesis was primed with an oligo-dT-XhoI primer and synthesized using M-MV reverse transcriptase. Second strand synthesis was accomplished with RNase H and T4 DNA polymerase. The double stranded cDNA was ligated to EcoRI linkers, digested with EcoRI and XhoI, and ligated into the phagemid vector lambda ZAP II XR. After construction of this directional library by Clontech, it was returned to the QIMR. During analysis of the library at the QIMR, we have found that a small percentage, 2% to 3%, of the clones contain inserts that appear to be highly homologous to sequences from salmonid fishes, as determined by homology comparisons using BLAST and by Southern hybridization analysis to genomic DNA from salmon (Sigma Chemical Co., St. Louis, MO) under stringent washing conditions. The remainder of the clones appear to contain S. japonicum sequences."

BASE COUNT 204 a 109 c 144 g 211 t
 ORIGIN

Query Match 58.0%; Score 636; DB 87; Length 668;
 Best Local Similarity 99.8%; Pred. No. 3.7e-172;
 Matches 647; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 8 CTATAGCTATTTGAAATTAAGGCGCTTGTCACACCGACCTCTTTTGGAAAT 67
 DB 1 CTATAGCTATTTGAAATTAAGGCGCTTGTCACACCGACCTCTTTTGGAAAT 60
 QY 68 ATCTTGAACAAAATATGAAGCATTTGTATGAGCGCATGAGTGATTAATGGCGAA 127
 DB 61 ATCTTGAACAAAATATGAAGCATTTGTATGAGCGCATGAGTGATTAATGGCGAA 120

QY 128 ACAAAGTTTGAAATGGCTTTGGAGTTTCCCAATCTCTTATATATGATGTGANG 187
 DB 121 ACAAAGTTTGAAATGGCTTTGGAGTTTCCCAATCTCTTATATATGATGTGANG 180
 QY 188 TTAATAATACAGCTATAGGCCATCATACGTTATATACCTGACAGCAACATGTGG 247
 DB 181 TTAATAATACAGCTATAGGCCATCATACGTTATATACCTGACAGCAACATGTGG 240
 QY 248 GTGCTGTCCAAAAGAGCGCTGCAGAGATTTCAATGCTTGAAGAGCGTTTGGATATTA 307
 DB 241 GTGCTGTCCAAAAGAGCGCTGCAGAGATTTCAATGCTTGAAGAGCGTTTGGATATTA 300
 QY 308 GATACGCTTTTCGAGATTTGATATATATGATTAAGACTTTGAACCTCAAGTTGATTTTC 367
 DB 301 GATACGCTTTTCGAGATTTGATATATATGATTAAGACTTTGAACCTCAAGTTGATTTTC 360
 QY 368 TTAAGCAAGCTACCTGAATGCTGAATGCTGAATGCTGAATGCTGAATGCTGAATGCT 427
 DB 361 TTAAGCAAGCTACCTGAATGCTGAATGCTGAATGCTGAATGCTGAATGCTGAATGCT 420
 QY 428 TTAATGCTATCATATGATTAACCATCTGACTCATGTTGATGAGCGCTTGATGTTT 487
 DB 421 TTAATGCTATCATATGATTAACCATCTGACTCATGTTGATGAGCGCTTGATGTTT 480
 QY 488 TATACATGACCCCAATGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 547
 DB 481 TATACATGACCCCAATGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 QY 548 TTGAACCTATCCCAAAATGATTAAGTACTTGAATTCAGCAAGTATATAGCATGGCTT 607
 DB 541 TTGAACCTATCCCAAAATGATTAAGTACTTGAATTCAGCAAGTATATAGCATGGCTT 600
 QY 608 TGCAGGCTGCGCAAGCGCTTGGTGGCGACCATCTCCCAAAAT 655
 DB 601 TGCAGGCTGCGCAAG-CACGTTTGGTGGCGACCATCTCCCAAAAT 647

RESULT 2
 A1932072 662 bp mRNA EST 20-MAR-2005
 LOCUS A1932072 Schistosoma japonicum cDNA 5', mRNA
 DEFINITION
 ACCESSION JAA000214.S3 Adult SJC 7/94 Schistosoma japonicum cDNA 5', mRNA
 VERSION A1932072
 KEYWORDS A1932072.1 GI:5670786
 SOURCE EST.
 ORGANISM Schistosoma japonicum.
 Schistosoma japonicum.
 Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea
 1 Strigeidae; Schistosomatidae; Schistosomatidae; Schistosoma.
 AUTHORS Hu, W., Brindley, P.J., and Feng, Z.
 TITLE Expressed sequence tags from adults of Schistosoma japonicum (Anhui strain) (Hu, Brindley, Feng)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Brindley, P.J.
 Molecular Parasitology Unit
 Queensland Institute of Medical Research
 300 Herston Road, Queensland 4029, Australia
 Tel: 61 7 3362 0413
 Fax: 61 7 3362 0104
 Email: paul.beglmr.edu.au
 PCR Primers
 FORWARD: M13 Forward
 BACKWARD: M13 Reverse
 Insert Length: 1200 Std Error: 0.00
 Seq primer: T3 Reverse
 High quality sequence stop: 549.
 Location/Qualifiers

FEATURES
 source 1..662
 /organism="Schistosoma japonicum"
 /strain="Chinese (Anhui) strain"
 /db_xref="taxon:6182"
 /clone_lib="Adult SJC 7/94"

```

/sex="Male and female"
/tissue_type="Whole body"
/dev_stage="Adult worms"
/lab_host="Mouse and rabbit"
/Note="Vector: Lambda ZAP-II XR; Site-1: EcoR I; Site-2:
XhoI I; Several hundred adult Schistosoma japonicum (Anhui
, P.R. China, strain), of mixed sex, were perfused from
the mesenteries of experimentally infected mice and
rabbits at the Queensland Institute of Medical Research,
Brisbane, Australia (QIMR), and stored for several months
in liquid nitrogen. Subsequently, mRNA was isolated at the
QIMR from lysates of these worms by oligo dt
chromatography, using a kit from Pharmacia. The mRNA was
then shipped to Clontech, Palo Alto, CA, USA, who
constructed a cDNA library. First strand synthesis was
primed with an oligo-dt-XhoI-primer and synthesized using
M-MuV reverse transcriptase. Second strand synthesis was
accomplished with RNase H and T4 DNA polymerase. The
double stranded cDNA was ligated to EcoRI linkers,
digested with EcoRI and XhoI, and ligated into the
phagemid vector lambda ZAP II XR. After construction of
this directional library by Clontech, it was returned to
the QIMR. During analysis of the library at the QIMR, we
have found that a small percentage, 2% to 3%, of the
clones contain inserts that appear to be highly homologous
to sequences from salmonid fishes, as determined by
homology comparisons using BLAST and by Southern
hybridization analysis to genomic DNA from salmon (Sigma
Chemical Co., St. Louis, MO) under stringent washing
conditions. The remainder of the clones appear to contain
S. japonicum sequences."
BASE COUNT      206 a      107 c      142 g      207 t
ORIGIN

```

```

Query Match      55.4%; Score 607.4; DB 27; Length 662;
Best Local Similarity 99.5%; Pred. No. 6.3e-164;
Matches 630; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

```

```

QY 23 GCAAAATTAGGCGCTTGCAACCCACTGACTCTTTTGGAAATCTTGAAGAAAT 82
    |||||||
Db 1  GGAAATTAAGGCGCTTGCAAC-CCACTGACTCTTTTGGAAATCTTGAAGAAAT 59
    |||||||
QY 83 ATGAAGACATTTGATGAGCGGATGAAGGATGAATGGCAACAAAGTTGAAT 142
    |||||||
Db 60 ATGAAGACATTTGATGAGCGGATGAAGGATGAATGGCAACAAAGTTGAAT 119
    |||||||
QY 143 TGGGTTTGGAGTTCCCAATCTTCTTATTAATGATGATGATTAATTAACACAGT 202
    |||||||
Db 120 TGGGTTTGGAGTTCCCAATCTTCTTATTAATGATGATGATTAATTAACACAGT 179
    |||||||
QY 203 CTATGGCCATCATAGCTTATAGCTGACAAACACATGTTGGTGTCCAAAG 262
    |||||||
Db 180 CTATGGCCATCATAGCTTATAGCTGACAAACACATGTTGGTGTCCAAAG 239
    |||||||
QY 263 AGCGTGCAAGATTTCAATGCTTGAAGAGCGGTTTGGATTAATAGATCGTTTGA 322
    |||||||
Db 240 AGCGTGCAAGATTTCAATGCTTGAAGAGCGGTTTGGATTAATAGATCGTTTGA 299
    |||||||
QY 323 GAATTCATATAGTAAGACTTTGAACCTCAAAAGTTGATTTCTTAGCAAGTACCTG 382
    |||||||
Db 300 GAATTCATATAGTAAGACTTTGAACCTCAAAAGTTGATTTCTTAGCAAGTACCTG 359
    |||||||
QY 383 AAATCTGAATAATGTTGAACATGTTTATGCTATTAACATATTTAAATGCTATCATG 442
    |||||||
Db 360 AAATCTGAATAATGTTGAACATGTTTATGCTATTAACATATTTAAATGCTATCATG 419
    |||||||
QY 443 TAACCATCTGACTTCAATGTTGATGAGCGCTTGAATGTTGTTTATACATGACCCAA 502
    |||||||
Db 420 TAACCATCTGACTTCAATGTTGATGAGCGCTTGAATGTTGTTTATACATGACCCAA 479
    |||||||
QY 503 TGTGCTGATGATGCTGCCAAATAGTTGTTTAAACGATATGAAGCTATCCAC 562
    |||||||

```

```

Db 480 TGTGCTGATGCTGCCAAATAGTTGTTTAAACGATATGAAGCTATCCAC 539
QY 563 AAATGATAGTACTTGAATCCAGCAAGTATATGACATGCGCTTGCAGGCTGGCAAG 622
    |||||||
Db 540 AAATGATAGTACTTGAATCCAGCAAGTATATGACATGCGCTTGCAGGCTGGCAAG 599
    |||||||
QY 623 CCAGCTTGGTGGTGGCGACCATCTCCAAAT 655
    |||||||
Db 600 -CACGTTTGGTGGTGGCGACCATCTCCAAAT 631

```

```

RESULT 3
AM186543
LOCUS
DEFINITION
JAA000443.SH3 Adult SjC 7/94 Schistosoma japonicum cDNA 5', mRNA
sequence.
VERSION
AM186543
ACCESSION
AM186543.1 GI:6459860
KEYWORDS
EST.
SOURCE
Schistosoma japonicum.
ORGANISM
Schistosoma japonicum
Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea
; Strigolidae; Schistosomatidae; Schistosomatidae; Schistosoma.
REFERENCE
1 (bases 1 to 601)
Hu, W., Brindley, P.J., and Feng, Z.
Expressed sequence tags from adults of Schistosoma japonicum (Anhui
strain) (Hu, Brindley, Feng)
Unpublished (1999)
Contact: Brindley, P.J.
Molecular Parasitology Unit
Queensland Institute of Medical Research
300 Herston Road, Queensland 4029, Australia
Tel: 61 7 3362 0413
Fax: 61 7 3362 0104
Email: paul@qimr.edu.au
PCR PRIMERS
FORWARD: M13 Forward
BACKWARD: M13 Reverse
Insert length: 1 Std Error: 0.00
Seq primer: T3 Reverse
High quality sequence stop: 601.
Location/Qualifiers
1. 601
/organism="Schistosoma japonicum"
/strain="Chinese (Anhui) strain"
/db_xref="taxon:6182"
/clone_lib="Adult SjC 7/94"
/sex="Male and female"
/tissue_type="Whole body"
/dev_stage="Adult worms"
/lab_host="Mouse and rabbit"
/Note="Vector: Lambda ZAP-II XR; Site-1: EcoR I; Site-2:
XhoI I; Several hundred adult Schistosoma japonicum (Anhui
, P.R. China, strain), of mixed sex, were perfused from
the mesenteries of experimentally infected mice and
rabbits at the Queensland Institute of Medical Research,
Brisbane, Australia (QIMR), and stored for several months
in liquid nitrogen. Subsequently, mRNA was isolated at the
QIMR from lysates of these worms by oligo dt
chromatography, using a kit from Pharmacia. The mRNA was
then shipped to Clontech, Palo Alto, CA, USA, who
constructed a cDNA library. First strand synthesis was
primed with an oligo-dt-XhoI-primer and synthesized using
M-MuV reverse transcriptase. Second strand synthesis was
accomplished with RNase H and T4 DNA polymerase. The
double stranded cDNA was ligated to EcoRI linkers,
digested with EcoRI and XhoI, and ligated into the
phagemid vector lambda ZAP II XR. After construction of
this directional library by Clontech, it was returned to
the QIMR. During analysis of the library at the QIMR, we
have found that a small percentage, 2% to 3%, of the
clones contain inserts that appear to be highly homologous
to sequences from salmonid fishes, as determined by.

```

1. 601
 /organism="Schistosoma japonicum"
 /strain="Chinese (Anhui) strain"
 /db_xref="taxon:6182"
 /clone_lib="Adult SjC 7/94"
 /sex="Male and female"
 /tissue_type="Whole body"
 /dev_stage="Adult worms"
 /lab_host="Mouse and rabbit"
 /Note="Vector: Lambda ZAP-II XR; Site-1: EcoR I; Site-2:
 XhoI I; Several hundred adult Schistosoma japonicum (Anhui
 , P.R. China, strain), of mixed sex, were perfused from
 the mesenteries of experimentally infected mice and
 rabbits at the Queensland Institute of Medical Research,
 Brisbane, Australia (QIMR), and stored for several months
 in liquid nitrogen. Subsequently, mRNA was isolated at the
 QIMR from lysates of these worms by oligo dt
 chromatography, using a kit from Pharmacia. The mRNA was
 then shipped to Clontech, Palo Alto, CA, USA, who
 constructed a cDNA library. First strand synthesis was
 primed with an oligo-dt-XhoI-primer and synthesized using
 M-MuV reverse transcriptase. Second strand synthesis was
 accomplished with RNase H and T4 DNA polymerase. The
 double stranded cDNA was ligated to EcoRI linkers,
 digested with EcoRI and XhoI, and ligated into the
 phagemid vector lambda ZAP II XR. After construction of
 this directional library by Clontech, it was returned to
 the QIMR. During analysis of the library at the QIMR, we
 have found that a small percentage, 2% to 3%, of the
 clones contain inserts that appear to be highly homologous
 to sequences from salmonid fishes, as determined by.


```

OY 481 GTTGTATATACATGAGCCCAATGTCCTG 510
Db 522 GTTGTATATACATGAGCCCAATGTCCTG 551

RESULT 5
AM017133
LOCUS AM017133 669 bp mRNA EST 10-SEP-1999
DEFINITION EST272555 Schistosoma mansoni male, Phil Loverde/Joe Merrick
ACCESSION Schistosoma mansoni cDNA clone SMAV95 5' end, mRNA sequence.
VERSION AM017133.1 GI:5869156
KEYWORDS EST.
SOURCE Schistosoma mansoni.
ORGANISM Schistosoma mansoni
Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea
REFERENCE 1 (bases 1 to 669)
AUTHORS Merrick,J.M., Osman,A., Loverde,P.T., Chandra,I., Glodek,A., Fraser
,C.M. and Lee,N.H.
TITLE Generation of a Schistosoma sp. EST (SMEST) Catalog & Schistosoma
Gene Index
JOURNAL Unpublished (1998)
COMMENT Contact: Norman H. Lee
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3529
Fax: 301 838 0208
Email: nhlee@ligr.org
Seq primer: M13 Reverse.

FEATURES
source
1..669
/organism="Schistosoma mansoni"
/db_xref="taxon:6183"
/clone="SMAV95"
/clone_lib="Schistosoma mansoni male, Phil Loverde/Joe
Merrick"
/sex="male"
/note="Vector: pBluescript SK; Site_1: EcoRI; Site_2: XhoI
; directionally cloned cDNAs"
BASE COUNT 216 a 114 c 139 g 199 t 1 others
ORIGIN

Query Match 40.6%; Score 445.2; DB 38; Length 669;
Best Local Similarity 81.3%; Pred. No. 3.1e-117;
Matches 516; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

OY 23 GGAATATTAAGGCGCTTGCAACCCGACGCTCTTTGGATATCTTGAAGAAAAT 82
Db 8 GGAAGTCAAGGCGCTTGACACACACGCTCTTTGGACACCTTGAGAACTT 67
OY 83 ATGAAGACATTTGATGAGCGCGATGAAGGATTAATGCGCAACAAAAGTTGAAT 142
Db 68 ATGAGAACGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 127
OY 143 TGGGTTGAGGTTCCCATCTCTCTTATTTATTTGATGATGATGATGATGATGATGAT 202
Db 128 TAGGCGCTGAGTCCCAATCTCTCTTATTTATTTGATGATGATGATGATGATGATGAT 187
OY 203 CTATGCGCATATACGTTATATAGCTGACAGCAACATGTTGGGTTGGTCCAAAG 262
Db 188 CTATGCGCATATACGTTATATAGCTGACAGCAACATGTTGGGTTGGTCCAAAG 247
OY 263 AGCGTGCAGATTTCAATGCTTGAAGAGCGGTTTGGATATTAGATACCGTGTTCGA 322
Db 248 AACGTCGGAATTTGATGCTTGAAGAGCGGTTTGGATATTAGATGATGATGATGATGAT 307
OY 323 GAATTCATTAAGTAAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAG 382
Db 308 GAATTCATTAAGTAAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAG 367

```

```

OY 383 AATGCTGAAGATTTGCAAGATCGTTATGTCATTAACATATTAAATGTCATAG 442
Db 368 GGGGCTGAAGATTTGCAAGATCGTTATGTCATTAACATATTAAATGTCATAG 427
OY 443 TAACCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 502
Db 428 TAACCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 487
OY 503 TGTGCTGATCGTTCGCCAATATAGTTGTTTAAAGATGATGATGATGATGATGATGAT 562
Db 488 AGTGTGTAAGAGTTTCCAAATATAGTTTCTTCAAAAAGTATTAAGATTACAC 547
OY 563 AATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 622
Db 548 AATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 607
OY 623 CCAGCTTTGTTGTTGCGACATCTCCCAATATG 657
Db 608 CTACGTTGTTGTTGCGACATCTCCCAATATG 642

RESULT 6
AM017155
LOCUS AM017155 636 bp mRNA EST 10-SEP-1999
DEFINITION EST272577 Schistosoma mansoni male, Phil Loverde/Joe Merrick
ACCESSION Schistosoma mansoni cDNA clone SMAV14 5' end, mRNA sequence.
VERSION AM017155.1 GI:5869178
KEYWORDS EST.
SOURCE Schistosoma mansoni.
ORGANISM Schistosoma mansoni
Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea
REFERENCE 1 (bases 1 to 636)
AUTHORS Merrick,J.M., Osman,A., Loverde,P.T., Chandra,I., Glodek,A., Fraser
,C.M. and Lee,N.H.
TITLE Generation of a Schistosoma sp. EST (SMEST) Catalog & Schistosoma
Gene Index
JOURNAL Unpublished (1998)
COMMENT Contact: Norman H. Lee
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3529
Fax: 301 838 0208
Email: nhlee@ligr.org
Seq primer: M13 Reverse.

FEATURES
source
1..636
/organism="Schistosoma mansoni"
/db_xref="taxon:6183"
/clone="SMAV14"
/clone_lib="Schistosoma mansoni male, Phil Loverde/Joe
Merrick"
/sex="male"
/note="Vector: pBluescript SK; Site_1: EcoRI; Site_2: XhoI
; directionally cloned cDNAs"
BASE COUNT 203 a 107 c 124 g 202 t
ORIGIN

Query Match 38.4%; Score 420.6; DB 38; Length 636;
Best Local Similarity 82.3%; Pred. No. 3.7e-110;
Matches 483; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

OY 69 TCTTGAAGAAATATGAGAGCATTTGATGAGCGCGATGAAGTGATTAATGCGGAAA 128
Db 8 TCTTGGCGAAGCATATGAGAGCGGTTGATGATGATGATGATGATGATGATGATGATGAT 67
OY 129 CAAAAGTTGAATGGGTTTGGAGTTTCCCAATCTCTTATTTATTTATGATGATGATGAT 188
Db 68 CGAAAGTTAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 127
OY 189 TAATTTAACAAGCTATATGCGCATCATACGTTATATAGCTGACAGCAACACATGTTGG 248

```

```

Db 128 TAAATTAACACATCTATGGCTATTTCATTAATAGCTGACAAACATTAATGTGGG 187
Oy 249 TGGTGTCCAAAGAGCGCGAGAGATTTCAATGCTTGAGAGCGCGTTTGGATATTAG 308
Db 188 TGGTGTCTCTAAGAGAGCGCGAATTTCTATGCTTGAGGAGCATTTTAGATATTAG 247
Oy 309 ATAGCGTGTCTTGAGAAATTCATATATATAGAAAGCTTTGAACCTCTCAAGTTGATTTCT 368
Db 248 ATATGCTGTCTTCAAGAAATTCATATATATAGAAAGCTTTGAACCTCTCAAGTTGATTTCT 307
Oy 369 TAGCAGCTACTGAAATGCTGAAATGCTGAAATGCTGAAATGCTGAAATGCTGAAATGCT 428
Db 308 CAACCAACCTCTGAAATGCTGAAATGCTGAAATGCTGAAATGCTGAAATGCTGAAATGCT 367
Oy 429 AATAGTGTATCATATGACCATCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 488
Db 368 GACGGGTGATTAAGTACATCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCT 427
Oy 489 ATACATGAGACCAATGTCCTGATGCGTCCCAAAATTAATGTTTAAACAAAGTAT 548
Db 428 ATACATGAGACCAATGTTTGGATTAATTTCCAAACAAATTAATGCTTCAACAAAGTAT 487
Oy 549 TGAACCTATCCCAAAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 608
Db 488 AGAAATTTTACCAACCAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 547
Oy 609 GCAGGAGTGGCAAGCGTTGTTGGTGGGAGCAATGCTTCCCAAAAT 655
Db 548 ACAAGTGTGTCGCCACTTTTGTGTGGTGGAGATGCTTCCCAAAAT 594

```

```

RESULT 7
LOCUS A1976995 475 bp mRNA EST 27-AUG-1999
DEFINITION EST2721589 Schistosoma mansoni male, Phil Loverde/Joie Merrick.
Schistosoma mansoni cDNA SMMAE18 5' end, mRNA sequence.
ACCESSION A1976995
VERSION A1976995.1 GI:5790167
KEYWORDS EST.
SOURCE Schistosoma mansoni.
ORGANISM Schistosoma mansoni.
REFERENCE Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea
AUTHORS ; Strigeldida; Schistosomatidae; Schistosomatidae; Schistosoma.
Merrick, J.M., Osman, A., Loverde, P.T., Chandra, I., Glodex, A., Fraser
, C.M. and Lee, N.H.
TITLE Generation of a Schistosoma sp. EST (SMEST) Catalog & Schistosoma
JOURNAL Gene Index
COMMENT Unpublished (1998)
Contact: Norman H. Lee
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3529
Fax: 301 838 0208
Email: nhlee@igf.org
Seq primer: M13 Reverse.
FEATURES
Source 1. 475
Location/Qualifiers
/organism="Schistosoma mansoni"
/db_xref="taxon:6183"
/clone="SMMAE18"
/clone_lib="Schistosoma mansoni male, Phil Loverde/Joie
Merrick"
/sex="male"
/note="Vector: pBluescript SK; Site_1: EcoRI; Site_2: XhoI
; directionally cloned cDNAs"
BASE COUNT 145 a 76 c 100 g 154 t
ORIGIN
Query Match 33.1%; Score 363; DB 27; Length 475;
Best Local Similarity 86.1%; Pred. No. 1.3e-93;

```

```

Matches 402; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
Oy 12 ACTAGTTATTTGAAAATTAAGGCGCTTGCAACCCACTGCACTTTTGAATATCT 71
Db 9 ACTGGTTATTTGAAAATTAAGGCGCTTGCAACCCACTGCACTTTTGAATATCT 68
Oy 72 TGAAGAAAATTAAGGCGCTTGCAACCCACTGCACTTTTGAATATCT 131
Db 69 TGAAGAAAATTAAGGCGCTTGCAACCCACTGCACTTTTGAATATCT 128
Oy 132 AAGTTTGAATTTGCGGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 191
Db 129 AAGTTTGAATTTGCGGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 188
Oy 192 AATTAACACAGTGTATGAGCGCATATGATTAATGATTAATGATTAATGATTAATGATTAAT 251
Db 189 AATTAACACAGTGTATGAGCGCATATGATTAATGATTAATGATTAATGATTAATGATTAAT 248
Oy 252 TTGTCCAAAGAGCGTGCAGAGATTTCAATGCTTGAAGAGCGGTTTGGATTAATGATA 311
Db 249 TTGTCCAAAGAGCGTGCAGAGATTTCAATGCTTGAAGAGCGGTTTGGATTAATGATA 308
Oy 312 CGGTGTTTGAAGATTTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 371
Db 309 TGGTGTTCAGAAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 368
Oy 372 CAAGTACCTGAAATGCTGAAATGCTGAAATGCTGAAATGCTGAAATGCTGAAATGCTGAA 431
Db 369 CCAACTCCCTGCAATGCTGAAATGCTGAAATGCTGAAATGCTGAAATGCTGAAATGCTGAA 428
Oy 432 TGGTGTTCAGAAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 478
Db 429 CGGTGTTTGAAGATTTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 475

```

```

RESULT 8
LOCUS AM329883 421 bp mRNA EST 28-JAN-2000
DEFINITION JAYL0240.GYL Schistosoma japonicum Lambda gt11 Express library
Schistosoma japonicum cDNA clone JAYL0240.GY 5', mRNA sequence.
ACCESSION AM329883
VERSION AM329883.1 GI:6806941
KEYWORDS EST.
SOURCE Schistosoma japonicum.
ORGANISM Schistosoma japonicum.
REFERENCE Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea
AUTHORS ; Strigeldida; Schistosomatidae; Schistosomatidae; Schistosoma.
Li, Y., Wu, Z.D., and Yu, X.B.
TITLE Expressed sequence tags from adults of Schistosoma japonicum
(Japanese strain) (Li, Y.; Wu, Z.D.; Yu, X.B.)
JOURNAL Unpublished (1999)
COMMENT Contact: Wu ZD
Department of Parasitology
Sun-Yat-sen University of Medical Sciences
Box 510089, 74# Zhongshan Er Road, Guangzhou, Guangdong, P.R.China
Tel: 86-20-87330566
Fax: 86-20-87331679
Email: zdwu62@163.net
FEATURES
Source 1. 421
Location/Qualifiers
/organism="Schistosoma japonicum"
/strain="Chinese"
/db_xref="taxon:6182"
/clone="JAYL0240.GY"
/clone_lib="Schistosoma japonicum Lambda gt11 Express
library"
/sex="Mix"

```

/note="Vector: Lambda g11 Sfi-Not; Site_1: EcoRI; Site_2: NotI; Several hundred adult Schistosoma japonicum(jiangxi, P.R.China, strain), of mixed sex, were perfused from the mesenteries of experimentally infected rabbits. Double-strain cDNA synthesized with the mRNA isolated from adult worm, was inserted into the bacteriophage lambda g11 Sfi-Not arms between EcoRI and NotI site of the lacZ gene. The cDNA library was constructed by Chen S.Z. at Nanjing Medical University, Nanjing, Jiangsu, P.R. China. (see: Chen Shuzhen, et al. Chinese Journal of Zoonoses 1997,13(6): 23-25)"

BASE COUNT 145 a 70 c 75 g 131 t
ORIGIN

Query Match 30.1%; Score 329.4; DB 89; Length 421;
Best Local Similarity 99.7%; Pred. No. 5.8e-84;
Matches 330; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 335 ATTCGATATAGTAAGACTTGAACCTCAAGTTGATTTCTTACGAGCTACCTGAA 384
DB 1 ATTCGATATAGTAAGACTTGAACCTCAAGTTGATTTCTTACGAGCTACCTGAA 60
QY 385 ATTCGAAATCTTGGAGATCGTTATGTCATFAAACAATTTAAATGTCATCTGA 444
DB 61 ATTCGAAATCTTGGAGATCGTTATGTCATFAAACAATTTAAATGTCATCTGA 120
QY 445 ACCATCTGATCTGATGATGAGCTCTGATGTTGTTTATACATGAGACCCAAAG 504
DB 121 ACCATCTGATCTGATGATGAGCTCTGATGTTGTTTATACATGAGACCCAAAG 180
QY 505 TGCCGATGCGTTCGCAAAATTAAGTTTAAAAAAGCATTAAGCATATCCACAA 564
DB 181 TGCCGATGCGTTCGCAAAATTAAGTTTAAAAAAGCATTAAGCATATCCACAA 240
QY 565 ATTCGATGATCTGATTAATCAAGATATAGATGCGCTTGCAGGCTGGCAACC 624
DB 241 ATTCGATGATCTGATTAATCAAGATATAGATGCGCTTGCAGGCTGGCAACC 300
QY 625 ACGTTGGTGGTGGCGACCATCTCCAAAT 655
DB 301 ACGTTGGTGGTGGCGACCATCTCCAAAT 331

RESULT 9
A1068236/c 368 bp mRNA EST 03-AUG-1998
LOCUS A1068236 Schistosoma mansoni, Phil Loverde/Joe Merrick Schistosoma
DEFINITION S-transferrase, mRNA sequence.
ACCESSION A1068236
VERSION A1068236.1 GI:3386203
KEYWORDS EST.
SOURCE Schistosoma mansoni.
ORGANISM Schistosoma mansoni.
REFERENCE Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea
AUTHORS 1 (bases 1 to 368)
Merrick,J.M., Osman,A., Loverde,P.T., Chandra,I., Glodok,A., Fraser
C.M. and Lee,N.H.
TITLE Generation of a Schistosoma sp. EST (SMEST) Catalog & Schistosoma
JOURNAL Unpublished (1998)
COMMENT Contact: Merrick, J.
State University of New York at Buffalo
Department of Microbiology, School of Medicine and Biomedical
Sciences, Buffalo, NY 14214, USA
Tel: (716)-829-2158
Fax: (716)-829-2177
Email: merrick@acsu.buffalo.edu
Seq primer: M13 Reverse.
location/Qualifiers
1.368

/organism="Schistosoma mansoni"
/db_xref="taxon:6183"
/clone="SMC053"
/clone_1db="Schistosoma mansoni, Phil Loverde/Joe Merrick"
/note="Vector: pBluescript SK; Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 118 a 79 c 60 g 111 t
ORIGIN

Query Match 20.9%; Score 229.6; DB 15; Length 368;
Best Local Similarity 78.1%; Pred. No. 3e-55;
Matches 289; Conservative 0; Mismatches 79; Indels 2; Gaps 1;

QY 281 TGCTGAGGAGCGGCTTTGGATATATAGATCGGTGTCAGATTCATATGTAAG 340
DB 368 TGCTGAGGAGCGGCTTTGGATATATAGATCGGTGTCAGATTCATATGTAAG 311
QY 341 ACTTGAACCTCTCAAAAGTATTTCTTACGAGCTCAAGATGCTGAATGCTGAATGCTG 400
DB 310 AATATGAACCTCTCAAAAGTATTTCTTACGAGCTCAAGATGCTGAATGCTGAATGCTG 251
QY 401 AAGATCGTTATGTCATTAACATATTTAAATGTCATGATTAACCATCTGACTTCA 460
DB 250 AAGATCGTTATGTCATTAACATATTTAAATGTCATGATTAACCATCTGACTTCA 191
QY 461 TGTTGATGAGCGCTCTGATGTTGTTTATACATGAGACCCAAATGCTGATCGCTTC 520
DB 190 TGTTGATGAGCGCTCTGATGTTGTTTATACATGAGACCCAAATGCTGATCGCTTC 131
QY 521 CAAATATAGTTGTTTAAAAAAGCATTAAGTGAAGTATCCCAATGATTAAGTATGTA 580
DB 130 CAAATATAGTTGTTTAAAAAAGCATTAAGTGAAGTATTAACCAATTAAGATCAACTTAA 71
QY 581 AATCAGCATATATATGATGATGCTCTTGCAGGCTGCGCAAGCAGCTTGGTGGCG 640
DB 70 AATCAGCATATATATGATGATGCTCTTGCAGGCTGCGCAAGCAGCTTGGTGGCG 11
QY 641 ACCATCTGCC 650
DB 10 ATACTCTCC 1

RESULT 10
AA999498/c 438 bp mRNA EST 28-SEP-1999
LOCUS AA999498 Schistosoma mansoni, cercariae Schistosoma mansoni
DEFINITION MCB3400.MIF Schistosoma mansoni, cercariae Schistosoma mansoni
CDNA clone SMC3400 3' end similar to glutathione S-transferase,
mRNA sequence.
ACCESSION AA999498
VERSION AA999498.1 GI:3190053
KEYWORDS EST.
SOURCE Schistosoma mansoni.
ORGANISM Schistosoma mansoni.
REFERENCE Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea
AUTHORS 1 (bases 1 to 438)
Santos,T.M., Johnston,D.A., Azevedo,V., Ridgers,I.L., Martinez,M.F.,
Marotta,G.B., Santos,R.L., Fonseca,S.J., Ortega,J.M., Rebelo
,D. and Pena,S.D.J.
TITLE Analysis of the gene expression profile of Schistosoma mansoni
JOURNAL M.O. Biochem. Parasitol. 103 (1), 79-97 (1999)
COMMENT Analysis using the expressed sequence tag approach
CONTACT: Santos, T.M. and Pena, S.D.J.
LABORATORY: de Genetica-Bioquímica, Departamento de Bioquímica e
Imunologia
INSTITUTO: de Ciencias Biológicas, Universidade Federal de Minas
Gerais
AVENIDA: Antonio Carlos 6627, Belo Horizonte, MG, Brazil, 31270-010
TEL: (5531)4415611
FAX: (5531)4415409
EMAIL: santos@mno.icb.ufmg.br


```

VERSION      AI446859.1  GI:6651649
KEYWORDS     EST.
SOURCE       Echinostoma paraensei.
ORGANISM     Echinostoma paraensei.
              Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea
              ; Echinostomida; Echinostomata; Echinostomatidae; Echinostoma.
REFERENCE    1 (bases 1 to 828)
              Adema,C.M., Leonard,P.M., DeJong,R.J., Day,H.L., Edwards,D.J.,
              Burell,G., Hertel,L.A. and Loker,E.S.
              Analysis of messages expressed by Echinostoma paraensei miracidia
              and sporocysts, obtained by random EST sequencing
              J. Parasitol. 86 (1), 60-65 (2000)
JOURNAL      20163500
MEDLINE
COMMENT      Contact: Loker ES
              Parasitology
              Biology, University of New Mexico
              Castetter Hall, Albuquerque, NM 87131, USA
              Tel: 505 277 5508
              Fax: 505 277 0304
              Email: esloker@unm.edu
              Insert length: 828   Std Error: 0.00
              Seq primer: T3 and T7.
FEATURES
  source      1. 828
              Location/Qualifiers
              /organism="Echinostoma paraensei"
              /strain="New Mexico, Loker"
              /db_xref="taxon:48215"
              /clone_lib="EPMs Lambda ZAP Express"
              /sex="hermaphrodite"
              /dev_stage="miracidia and sporocysts"
              /note="Vector: Lambda ZAP Express; Site_1: Xho I; Site_2:
              EcoRI"
BASE COUNT   256 a 165 c 169 g 238 t
ORIGIN
Query Match          20.0%; Score 219.6; DB 20; Length 828;
Best Local Similarity 58.6%; Pred. No. 3e-52;
Matches 381; Conservative 0; Mismatches 269; Indels 0; Gaps 0;
OY 1 ATGTCCCATATCTAGTATGTAAGAAATTAAGGCGCTTGCAACCCCGCTGACTCTT 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 62 ATGGCTCTGTGTAGATATGGCTCTCGCTGGACGATGCCAACCAATTTGTCTTTT 121
OY 61 TTGGAATATCTTGAAGAAAATATGAAGAGCATTTGTATAGCGGAGTGAAGATATAA 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 122 CTGGATATCTTGAAGAACCATACGAGAAAGCGTATGATGATGCGGGA 181
OY 121 TGGCGAACAATAAGTTGANTGGGTTGGAGTTTCCCATCTCTTATATATGAT 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 182 TGGCAAGCGGAGAACCACTCCATTAAGCTGATTTTCCAAACCTTCTATTACATGAT 241
OY 181 GGTGATGTTAAATTACACAGTCTATGGCCATCATACCTTATATAGTGACACACAAC 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 242 GGTGACTTTTGATATACAGTCTTGCCATTTATCAGTATCTTCCGATTAAGACAAAT 301
OY 241 ATGTGGGTGTTGTCACAAAGCGTGCAGAGATTTCAATGCTTGAAGAGCGGTTTGG 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 302 CTCTGGGTGACGACCCCAAGAGACGTCCTATATACATATGCTGAGAGTGGCGTACT 361
OY 301 GATATTATAGGTTGTTGAGATTCATATAGTAAGCTTGAAGCTTGAACCTCAAAAT 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 362 GATTTAGCATGGCAATTCGCAACAATGCTTACAGTCTCTCAGATGAGAAAGAGAAAGCT 421
OY 361 GATTTTCTTACCAAGCTACCTGAAATGCTGAAATGTTGCAAGATCTTTTATGTCATPAA 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 422 AGTGTATCGAAGAACTCCGACGCTGAAAGAGTTGCAAAAGTATTTGGAATCGAAA 481
OY 421 ACATATTTAAATGATGATGTAACCATCTCGACCTTCATGTTGATAGACCTCTTGAT 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 482 CCAATTCCTAAAGGAGACAAACCACTGATCTGATCTCATGTTTACGAAACACTGAT 541
OY 481 GTTGTATTATACATGAGCAACATGTGCTGATGCTGCCAACAAATTTGTTTAA 540

```

```

DB 542 TGATCAACTAGTGGATTCMAAACATTAGATTCATTTCTTACGTGGAAGACTACATA 601
OY 541 AAACGATTGAAGCATATCCCAAAATGATGATGATGTAAGTGAATTCAGCAAGTATPAGA 600
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 602 AACAGATGAGAAATTTACCCCAATCAAAAGCATATATGATGATGATGATTTGCAAA 661
OY 601 TGCGCTTTCAGGCGTGGCAAGCCAGTTGGTGGTGGGACCATCTTC 650
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 662 TGCGCACTGCAATGCTGTTTCTTCTGCAATTTGGTGGGAGCAACCACTTC 711
RESULT 13
AM654220 550 bp mRNA EST 14-JUL-2000
LOCUS     AM654220
DEFINITION 103527 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION AM654220
VERSION   AM654220.1 GI:7420046
KEYWORDS  EST.
SOURCE    EST.
ORGANISM  Bos taurus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
          Bovidae; Bovinae; Bos.
          1 (bases 1 to 550)
          Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M.,
          Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Laegreid
          ,W.W. and Keele,J.W.
          Design and use of four pooled tissue normalized cDNA libraries for
          EST discovery in cattle
          Unpublished (2000)
          Contact: Smith TPL
          USDA, ARS, US Meat Animal Research Center
          PO Box 166, Clay Center, NE 68933-0166, USA
          Tel: 402 762 4366
          Fax: 402 762 4390
          Email: smith@email.marc.usda.gov
          Single pass sequencing. Bases called and alt trimmed with phred
          v0.980904.e. Vector identified by cross_match with the -minscore 18
          and -mismatch 12 options.
          PCR primers
          FORWARD: AGGAACAGCTAGACCAT
          BACKWARD: GTTTCCTCAGTCAGCAG
          Plate: 94 row: F column: 6
          Seq primer: ATTATGCTGACACTATAG.
FEATURES
  source      1. 550
              Location/Qualifiers
              /organism="Bos taurus"
              /db_xref="taxon:9913"
              /clone_lib="MARC 1BOV"
              /tissue_type="pooled"
              /lab_host="DH10B"
              /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
              Library made from pooled tissue from lymph node, ovary,
              fat, hypothalamus, and pituitary."
BASE COUNT   135 a 142 c 148 g 125 t
ORIGIN
Query Match          11.7%; Score 127.8; DB 93; Length 550;
Best Local Similarity 54.2%; Pred. No. 6.9e-26;
Matches 292; Conservative 0; Mismatches 232; Indels 15; Gaps 1;
OY 1 ATGTCCCTATCTAGTATTTGAAATTAAGGCGCTTGCAACCCGCTGACTCTT 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7 ATGCCCATGATCTCGGTTACTGGGACATCCGCGGCTGGCCCTATTCGCTTCTC 66
OY 61 TTGGAATATCTTGAAGAAAATATGAAGAGCATTTGTATGAGCGCG----- 106
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 67 CTGAGATACACAGACAACTATAGAGAGAGGACGATACGCTGATGAGATGCTCCGAC 126
OY 107 -ATGAAGTATTAATGCGCAACAAAGTTGATTTGGGTTTGAAGTTTCCCATCTT 165
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


Db	127	TATGACGAAGCCAGCTGGCTGAATGAAAAAATCAAGCTGGGCGCTGGACCTTCCCAATCTG	186
Qy	166	CCTTATTATATTGATGTGATGTTAAATTAAACACAGCTCTATGGCATCATAGTTATATA	225
Db	187	CCCTACTTAATCGATGGGACTGCACAACTCACCCAGAGCAACGCCATCTCTCGTACATC	246
Qy	226	GCTGACAGGACACATGTTGGGGGTTCGCAAAAAGAGGCTGCAGAGATTCAATGCTT	285
Db	247	GCTGCGACAGCAACCTGTGTGGGGAGACAGAGAGAGATGATTGCTGTGGACATTTTGG	306
Qy	286	GAAGAGACGGTTTTGATATTAGATACGGTGTTCGAGATTGCTCATAGTAAAGACTTT	345
Db	307	GAGAACACAGGTTATAGGATGTCCGTTGGCCATGGCTAGGATCTGCTACAGCCCTGACTTT	366
Qy	346	GAAACTCTCAAAATGTAATTTCTTAGCAAGCTACCTGAAATGCTGAAATGTTGGAAGAT	405
Db	367	GAGAAACTGAAGCCTCGTTTCTTGAAGAGATCCCTGAAAAAATCAACGTGTTCTCAGAC	426
Qy	406	CGTTTATGTCTATAAACATATTTTAAATGATGATCATGTAAACCATCTGACTTCATGTTG	465
Db	427	TTTTCTGGGGAAGAGCCCTTGTTGGCAGGGGACAAAGCTCACATGTGATTTCTCGTTT	486
Qy	466	TATGACCTCTGATGTGTTTTTATACATGAGACCAATGTGCTGGATGGCTCCCAAA	524
Db	487	TATACGCTCTTGACATGACACCGCATATTGTAGAGCCCAAGTGCCCTGGAGGCATTTCCAAA	545

LOCUS	DEFINITION	AM239463	624 bp	mrna	EST	13-DEC-1999
AM239463	LOCUS	AM239463	624 bp	mrna	EST	13-DEC-1999
DEFINITION	similar to gb:m63509 GLUTATHIONE S-TRANSFERASE MUSCLE (HUMAN);, mRNA sequence.	x440c08.v1 NCI CGAP Lu31 Homo sapiens cDNA clone IMAGE:2578766	5			

ACCESSION	AW239463
VERSION	AW239463.1
KEYWORDS	EST.
SOURCE	human.

ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 624)
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550
 Email: Robert_Strausberg@nih.gov
 Tissue Procurement: ATCC CDNA Library Preparation: Life
 Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The
 I.M.A.G.E. Consortium DNA Sequencing by: Washington University
 Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdipr/image/image.html
 Seq primer: -40RP from Gibco
 High quality sequence stop: 435.
 Location/Qualifiers
 1..624

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:257876"
/clone_id="NCI_CGAP_Lu31"
/sex="male"
/dev_stage="fetal, 14 wk post-conception"
/lab_host="DH10B"
/note="Organ: lung, cell line: PCMW-SPORT6;
Site_1: EcorV; Site_2: NotI; Cloned unidirectionally, no
5' adaptor. Primer: Oligo dT. Full-length library
constructed by life technologies."
BASE COUNT      166 a      166 c      154 g      138 t
ORIGIN

```

Query Match	11.6%;	Score 127;	DB 88;	Length 624;	.
Best Local Similarity	51.8%;	Pred. No. 1.2e-25;			
Matches 323; Conservative	0;	Mismatches 285;	Indels 15;	Gaps 1;	

OY	1	ATGTCCTTACTAGGTATTATGGAAATAA	TAAAGGCCCTTGTCACCCACTGCATCTT	60
Dd				
Dd	2	ATGCCCATGACACTGGGGTACTGGACATC	GCGCGGGCTGGCCCATTTCCATCCGCGCTC	61
OY	61	TTGGAATATCTTGAAGAANAATATGACA	GACCATTTGTATGAGCCCATGGAAGTGATAA	120
Dd				
Dd	62	CTGGAATATCACAGACTCAAGCTACGAG	AAAAAGAAATGACAGATGGGGAGCGCTCGAT	121
OY	121	T-----	-GGCGAAACAAAAGTTTAATGGTGGTTG	AGTTTCCCAATCTT 165
Dd	122	TATGACAGAAAGCCGATGGCTGAATGA	AAAAAATTCAAAGCTGGGCCCTCGACTTT	CCCAATCTG 181
OY				
OY	166	CCTTATATATATGATGGTGAATGTTAA	ATTAACTAACACAGTCTATGGCCATCATAC	GTTATATA 225
Dd				
Dd	182	CCCCTACTGATGATGGAGACTCACAAAT	ACCACGAGACAAGCCATCTCGCGGTACATT	241
OY	226	GCTGACACGACACATGTTGGGTGGTGT	GTTCACAAAGAGCGTGAGAGATTTCCAACTT	285
Dd				
Dd	242	GCCCGCAAGCACACACTGTGGGGGAT	CANAAAAAGACAGATTCGCCAAGACATTTG	301
OY				
OY	286	GAAAGAGCGGTTTGGATATTAGATPAC	GGTGTTCGAAATTTGCATATAGTAAGACTTT	345
Dd				
Dd	302	GAGAACCCAGTTTATGGACAGCGCTAT	GCAAGCTGGCCCAAACCTCTGTATGACCC	GATTTT 361
OY	346	GAACCTCTCAAAAGTTGATTTTCTTAG	CAGACTTACCTTAATATGCTGAANAATGTT	GAAGAT 405
Dd				
Dd	362	GAGAACTCGAAACACAGATNACTGCAG	BGCACTCCCTGAATGCTGAAGCTCTACTC	AAG 421
OY	406	CGTTTATGTCTAATAACATVATTAAT	GTGATCATATGTAACCATCTGCACTTCAT	GTG 465
Dd				
Dd	422	TTTTCTGGGGAAGCGCCATGGTCTTCT	TGGGGACAAAGATCACTTTGGGATTTTCAT	GTCT 481
OY	466	TATGACGCTCTTGATNGTTGTTTATP	ATCATGGACCAACATNGGCCGATGGTTCC	CAAAA 525
Dd				
Dd	482	TATGATGTCTTGAGAGAAACAGATATT	TGAGGCCAGCTDCCCTGGATGCTTTCC	CAAC 541
OY	526	TTAGTTGTTTTTAAAAAACGTAATTGA	AGCTATCCACAAAATGATAAGTGAANAATC	585
Dd				
Dd	542	CTGAAGGACTCATCTCCGCTTTGAG	GGGCTTGAGAAAGATCTGTGCTCATGAT	GTCC 601
OY	586	AGCAAGTATATAGCATGGCCTTT	608 :	
Dd				
Dd	602	AGGCGCTTCCTCCCAAGACCTGT	624	

RESULT	15
LOCUS	AV590629
DEFINITION	AV590629 Bos taurus brain fetus Bos taurus cDNA clone EIBR013G03
ACCESSION	AV590629
VERSION	AV590629.1 GI:9701622
KEYWORDS	EST.
SOURCE	cst.
ORGANISM	Bos taurus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea Bovidae; Bovinae; Bos. 1 (bases 1 to 544)
AUTHORS	Sugimoto,Y., Hirotsune,S., Takasuga,A., Itoh,R., Jitohzono,A. and Suzuki,H.
TITLE	bovine cDNA sequencing
JOURNAL	Unpublished (2000)
COMMENT	Contact: yoshikazu sugimoto Animal Genetics Division Shirakawa Institute of Animal Genetics Ogakura, Nishigo, Nishi-Shirakawa, Fukushima 961-8061, Japan Tel.: 81-248-25-5641

Fax: 81-248-25-5725
Email: kazusugie@ccoca.ocu.ne.jp

Single pass sequencing.

This clone was obtained from a polyA-deleted cDNA library.

FEATURES

SOURCE

Location/Qualifiers
1..544

/organism="Bos taurus"

/db_xref="taxon:9913"

/clone="E1BR013G03"

/clone_1lb="Bos taurus brain fetus"

/tissue.type="brain"

/dev_stage="fetus"

/lab_host="DH10B"

/note="Vector: pZL1; site_1: SalI; site_2: NotI; Poly A
was deleted from a NotI site"

BASE COUNT 131 a 141 c 148 g 124 t

ORIGIN

Query Match

Best Local Similarity 10.9%; Score 119; DB 37; Length 544;
Matches 282; Conservative 0; Mismatches 230; Indels 15; Gaps 1;

```
QY 1 ATGCCCCCTATACAGGTATTTGAATAATTAGGCGCTTGCACCCACTGACTTCTT 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 18 ATGCCCATGATCCCTGGGTACTGAGACATCCGGGCTGCGCCATGCCATCCGCTTCTC 77
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 TTGGAAATATCTTGAAGAAAATATGAAGCATTTGTATGAGCGC----- 106
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 78 CTGAGATACACAGACACAAACATATGAGAGAGAGAGAGAGAGATGCTCCGAC 137
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 107 -ATGAAGTGATATAATGGCGAAACAAAGTTGAATGGGTTTGGAGTTTCCCAATCTT 165
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 138 TATGACAGAGCCAGTGGCTGAATGAATAATCAAGCTGGGCTGGACTCCCAATCTG 197
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 166 CCTTATTATATGATGATGATTTAATTAACAGAGTATAGGCGCATATAGCTTATATA 225
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 198 CCTTACTTATGATGATGAGCTACAGAGCTACACAGAGAGAGAGAGAGAGATTTTG 257
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 226 GCTGACAGACACATATGTTGGTGGTGTCCAAAAGAGCGTGCAGAGATTCAATGCTT 285
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 258 GCTGCGAAGCAGACAACTGTGTGGGAGACAGAGAGAGAGATGCTGTGACATTTTG 317
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 286 GAAGAGAGCGTTTGGATTTAGATAGCGTGTTCGAGATTCATATATGAAGACTTT 345
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 318 GAGAACAGAGTTATGATGATGCTGGCATGGCTAGAGATCTGACAGCCCTGACTTT 377
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 346 GAAACTCTCAAGTATTTCTTAGCAAGTACCTGAATGCTGAAATGTTGAAGAT 405
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 378 GAGAAACTGAGAGCGCTTGTTCAGAGGAGATCCCTGAAAAAATCAAGCTGTTCTCAG 437
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 406 CGTTTATGTCATATAAATATTTAAATGATGATGATTAACCATCCGACTTCATGTTG 465
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 438 TTTTCGGGGAGAGAGCGCTTGTTCAGAGGAGATCCCTATGATGATTTCTGCTT 497
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 466 TATGAGCGCTGTGATGTTGTTTATACATGAGCCCAATGCTGGA 512
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 498 TATGAGCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 544
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: March 18, 2001, 04:33:44
Job time: 433 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2001, 04:57:22 ; Search time 77.34 Seconds
(without alignments)
151.648 Million cell updates/sec

Title: US-09-402-488a-2
Perfect score: 1850
Sequence: 1 MSPILGYWKIKGLVQPIRL.....PKQSHNDGFEEIPEEYLIQ 343

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues
Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_36.*
1: /cgn2_2/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /cgn2_2/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /cgn2_2/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /cgn2_2/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /cgn2_2/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /cgn2_2/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /cgn2_2/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /cgn2_2/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /cgn2_2/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /cgn2_2/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /cgn2_2/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /cgn2_2/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /cgn2_2/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /cgn2_2/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /cgn2_2/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /cgn2_2/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /cgn2_2/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /cgn2_2/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /cgn2_2/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /cgn2_2/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /cgn2_2/gcgdata/geneseq/geneseq/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1850	100.0	343	20	W87636
2	1236	66.8	621	21	Y52289
3	1235	66.8	257	20	W74124
4	1235	66.8	282	20	W74123
5	1235	66.8	422	20	Y23648
6	1234.5	66.7	538	19	W47001
7	1234.5	66.7	564	16	R84565
8	1234.5	66.7	643	16	R84568
9	1234	66.7	282	20	W74096
10	1233.5	66.7	824	21	Y52288
11	1232.5	66.6	262	20	W74128
12	1232.5	66.6	282	20	W74127

13	1232	66.6	248	20	W74113	GST/GI transport r
14	1232	66.6	248	20	W74122	GST/GI transport r
15	1232	66.6	258	20	W74117	GST/GI transport r
16	1232	66.6	259	20	W74110	GST/GI transport r
17	1232	66.6	277	20	W74116	GST/GI transport r
18	1232	66.6	277	20	W74098	GST/GI transport r
19	1231	66.5	247	20	W74103	GST/GI transport r
20	1230.5	66.5	257	20	W74126	GST/GI transport r
21	1230	66.5	247	20	W74114	GST/GI transport r
22	1230	66.5	248	20	W74101	GST/GI transport r
23	1230	66.5	258	20	W74104	GST/GI transport r
24	1230	66.5	259	20	W74130	GST/GI transport r
25	1230	66.5	259	20	W74118	GST/GI transport r
26	1230	66.5	264	20	W74129	GST/GI transport r
27	1230	66.5	267	20	W74106	GST/GI transport r
28	1229.5	66.5	257	20	W74109	GST/GI transport r
29	1229.5	66.5	277	20	W74108	GST/GI transport r
30	1229	66.4	259	20	W74125	GST/GI transport r
31	1228	66.4	247	20	W74120	GST/GI transport r
32	1228	66.4	248	20	W74112	GST/GI transport r
33	1228	66.4	248	20	W74121	GST/GI transport r
34	1228	66.4	248	20	W74100	GST/GI transport r
35	1228	66.4	248	20	W74102	GST/GI transport r
36	1228	66.4	249	20	W74115	GST/GI transport r
37	1228	66.4	257	20	W74111	GST/GI transport r
38	1228	66.4	257	20	W74119	GST/GI transport r
39	1228	66.4	257	20	W74105	GST/GI transport r
40	1228	66.4	277	20	W74099	GST/GI transport r
41	1228	66.4	277	20	W74107	GST/GI transport r
42	1228	66.4	279	20	W74097	GST/GI transport r
43	1228	66.4	282	20	W74095	GST/GI transport r
44	1227	66.3	232	16	R72793	Glutathione-S-tran
45	1227	66.3	1196	21	Y58917	Fumonisin esterase

ALIGNMENTS

RESULT 1	
ID	W87636 standard; Protein: 343 AA.
XX	W87636;
AC	09-MAR-1999 (first entry)
XX	
DT	The fusion protein GST-chymosin pro-peptide-Hirudin.
XX	
DE	
XX	
KW	Fusion protein; bovine chymosin pro-peptide; leech; hirudin;
KW	anticoagulant protein; autochemically maturing zymogen;
KW	carp growth hormone; vaccine.
XX	
OS	Synthetic.
OS	Bos sp.
XX	Hirudo medicinalis.
XX	
FH	Key
FT	Peptide
FT	Location/Qualifiers
FT	/note="Bovine chymosin pro-peptide"
FT	237..278
FT	279..344
FT	/note="leech hirudin"
FT	263..264
FT	/note="pseudochymosin cleavage site"
XX	
PN	WO9849326-A1.
XX	
PD	05-NOV-1998.
XX	
PF	23-APR-1998; 98MO-CA00398.
XX	
PR	25-APR-1997; 97US-0044254.
XX	
XX	(SEMB-) SEMBIOSYS GENETICS INC.

XX Alcantara J, Moloney M, Van Rooijen G;
 FI WPI: 1999-059646/05.
 XX DR N-PSDB: V83966.
 XX
 PT Preparation of recombinant polypeptides - by expression of a fusion
 PT protein comprising a pro-peptide derived from an autocatalytically
 PT maturing zymogen and a heterologous polypeptide
 XX
 PS Claim 27; Fig 1; 44pp; English.

XX The present sequence represents a fusion protein comprising
 CC glutathione-S-transferase (GST)-bovine chymosin pro-peptide-leech
 CC hirudin. The chymosin pro-peptide sequence is placed upstream of the DNA
 CC sequence encoding the leech anticoagulant protein hirudin. The fusion
 CC protein was made to exemplify the invention. The specification describes
 CC a method for preparing a recombinant polypeptide in a host cell. A
 CC chimeric nucleic acid sequence encoding a fusion protein comprising
 CC a pro-peptide derived from an autocatalytically maturing zymogen linked
 CC to a protein heterologous to the pro-peptide, is introduced into the host
 CC cell. The host cells are then grown to produce the fusion protein.
 CC Altering the environment of the fusion protein allows cleavage of the
 CC pro-peptide from the fusion protein to release the recombinant
 CC polypeptide. The method can be used for the preparation of recombinant
 CC proteins can be used for delivering to a human or animal a therapeutic
 CC or nutritional polypeptide such as a vaccine, a peptide antibiotic, a
 CC cattle feed enzyme, a cytokine, a gastric lipase or a lactase.
 XX

Sequence 343 AA:

Query Match 100.0%; Score 1850; DB 20; Length 343;
 Best Local Similarity 100.0%; Pred. No. 1.7e-155;
 Matches 343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSPILGWKIKGLVOPTRLLLEYLEEKYEHLIERDEGDKMKRKKELGEPNLPYYTD 60
 Db 1 mspllgwkwikglvoptrlllleyleekyeehlyerdegdkwnkkkrlglfepnlpyytd 60
 QY 61 GGVKLTQSMATIRYIADKNNMGGCPKERAETISMLEGAVLDIRYGVSRVANSKDEFTLKV 120
 Db 61 ggvkltqsmatirytadknnm1ggcpkeraetismlegavldirvgvsrvtayskdftelkv 120
 QY 121 DFLSKLPKEMKMFEDRLCKRTYLNQGHVHPFMVYDALDVVLYMDPKCLAFKLVCKR 180
 Db 121 dflsklpemkmfedrlckrtlyngdvhlpdflmlydaldvvlmdpncldlfpklvckfk 180
 QY 181 KRLEAIPOIDKYLKSSKRYIAMPLOGMATFFGGDHPKSDLVPRGSPNSRVSSGRAEFT 240
 Db 181 krlaepoidkylksskrylamlplogmatffggdhpksdlvprgspnsrvdssgraeft 240
 QY 241 RIPIVKGKSLRKALKKEHGLEPFLQKQOYGISKSYSGFVYVYTDCTESGONLCEGSNVC 300
 Db 241 riipivkgkslrralkkkehglepflqkqoygisksygsfvvytdctesggnlcegsnvc 300
 QY 301 GCGNKCILGSDGKKNOCVWGECTPKPQSHNDGDEFEETBEFLQ 343
 Db 301 gcgncilgsgdggknocvwtgectpkpqsndgdfeetbeeylq 343

RESULT 2
 Y52289
 ID Y52289 standard; Protein; 621 AA.
 AC Y52289;
 XX
 XX
 DT 09-FEB-2000 (first entry)
 XX
 DE Human HCP phosphodomain/GST fusion protein.
 XX
 KW Haematopoietic cell phosphatase; HCP; cloning; expression;

KW epithelial cell; growth factor receptor; interleukin-3; IL-3;
 KW erythropoietin; EPO; negative regulator; signal transduction;
 KW modulation; activity; inhibitor; stimulation; cytokine therapy;
 KW antisense therapy; gene therapy; treatment; diagnosis; phosphodomain;
 KW catalytic domain; glutathione-S-transferase; fusion.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO954450-A1.
 XX
 PD 28-OCT-1999.
 XX
 PF 15-APR-1999; 99WO-US08228.
 XX
 PR 17-APR-1998; 98US-0082072.
 XX
 PA (ORTH) ORTHO-MCNEIL PHARM INC.
 XX
 PI Jolliffe L, Barbone F, Shanahan M, Xu D;
 DR WPI: 2000-013244/01.
 XX
 DR N-PSDB: 232917.

Assay methods for modulators of hematopoietic cell phosphatase
 Claim 3; Fig 4D; 57pp; English.

XX This sequence represents a human haematopoietic cell
 CC phosphatase (HCP) catalytic domain/glutathione-S-transferase (GST)
 CC fusion protein. This was cloned and expressed in E. coli and
 CC subsequently purified for use in a variety of activity assays.
 CC HCP is a 68 kD non-membrane associated protein found primarily in
 CC haematopoietic cells, and to some extent in epithelial cells. It
 CC has been shown to associate with several growth factor receptors
 CC of receptor activation, and hence of haematopoietic signal transduction.
 CC The receptors it regulates include the interleukin-3 (IL-3) and
 CC erythropoietin (Epo) receptors. The invention relates to a novel
 CC method of identifying compounds that modulate HCP activity. This
 CC comprises combining a modulator of HCP protein activity with HCP protein
 CC and HCP protein substrate, and measuring an effect of the modulator on
 CC the activity of the HCP protein. Inhibitors of HCP can be used to
 CC modulate the activity of haematopoietic receptors. HCP has been shown to
 CC be a negative regulator of one or more signal transduction pathways in
 CC haematopoietic cells, and the identification of a HCP inhibitor would
 CC provide a synthetic stimulator to the haematopoietic system could be
 CC used in conjunction with other cytokine therapy, e.g., administration of
 CC EPO. HCP is also associated with several growth factors found in
 CC haematopoietic cells and it is possible that a HCP inhibitor may also
 CC function as a haematopoietic cell potentiator, enhancing the effect of
 CC growth factors, decreasing the necessary dose of the factor. The HCP
 CC polynucleotides are sources of probes and primers, and can be used to
 CC design antisense sequences, and in gene therapy. The protein, or
 CC active fragments thereof, may be used to treat or diagnose disorders
 CC which require the modulation of the HCP activity. Small molecules that
 CC inhibit the action of HCP can also augment the effect of EPO, which
 CC meaning that a lower therapeutic dose of EPO may be utilised.

Sequence 621 AA:

Query Match 66.8%; Score 1236; DB 21; Length 621;
 Best Local Similarity 70.2%; Pred. No. 5.5e-101;
 Matches 257; Conservative 20; Mismatches 63; Indels 26; Gaps 7;

QY 1 MSPILGWKIKGLVOPTRLLLEYLEEKYEHLIERDEGDKMKRKKELGEPNLPYYTD 60
 Db 1 mspllgwkwikglvoptrlllleyleekyeehlyerdegdkwnkkkrlglfepnlpyytd 60
 QY 61 GGVKLTQSMATIRYIADKNNMGGCPKERAETISMLEGAVLDIRYGVSRVANSKDEFTLKV 120
 Db 61 ggvkltqsmatirytadknnm1ggcpkeraetismlegavldirvgvsrvtayskdftelkv 120

```

QY 121 DFLSKLPEMLKMFEDRLCHRTYLNCDHVTHPDFMLYDALDVLYMDPMCLDAPKLVCFK 180
DB 121 df1sklpemlkmedfclchktylngdhvchpdtmlydaldvlymdpmcldapklvcfk 180
QY 181 KRIEALPQIDKYLKSSKYIAMPLOQWQATFGGDPKPSDLVPRGSPN----- 228
DB 181 kriealpqidkylksskylawplqgwqatf9g9dhpksdlvprgspefgyfyllrpyy 240
QY 229 -SRVDSGSAEI-TRIPLKGSLSRKALKEHGLEDF--LQKQY-GISSKYSGFVYYD 283
DB 241 actvna---adlenrvlelnkkqesedtakagfweefsiqkqevknlhgrlegqrpenk 297
QY 284 CTESGONLCLCGSSNVCQGQ--NKCILGSD-----GEKNOCVWGEQTPKPSQSHNDGFEEI 337
DB 298 gknryknllpfhshrllqgrdsnlpsgdylnanylnqnlqgdenaktyiasqgcleat 357
QY 338 PERYLQ 343
DB 358 vndfwg 363

RESULT 3
W74124
W74124 standard; peptide: 257 AA.
W74124:
04-MAY-1999 (first entry)
GST/GI transport receptor binding protein fusion sequence.
Gastro-intestinal transport receptor; binding protein; hsi; HPT1;
D2H; hPEPT1; human; GI tract receptor; sucrose-isomaltase complex;
intestinal peptide-associated transporter; hypertension; diabetes;
osteoporosis; haemophilia; anaemia; cancer; migraine; angina pectoris;
therapeutic agent delivery; therapy; fusion protein.
Homo sapiens.
W09851325-A2.
19-NOV-1998.
15-MAY-1998; 98WO-US10088.
15-MAY-1997; 97US-0046595.
(CYTO-) CYTOGEN CORP.
(ELAN-) ELAN CORP PLC.
Alvarez VL, Belinka BA, Cagney GM, Carter JM, Lambkin IJ,
Omahony DJ, Patterson CA, Singleton J,
WPI: 1999-009568/01.
New proteins that bind specifically to receptors in the
gastro-intestinal tract and related nucleic acid - chimaeras and
antibodies, used to deliver therapeutic or diagnostic agents to, or
through, the gastrointestinal tract, e.g. insulin or leuprolide
Example: Page 224; 294pp; English.
This sequence represents a fusion protein between
glutathione-S-transferase and a gastro-intestinal transport protein
binding peptide. The invention relates to purified proteins (I) that bind
specifically to at least one of the gastro-intestinal (GI) tract
receptors human intestinal peptide-associated transporter (HPT1),
hPEPT1, D2H and human sucrose-isomaltase complex (hSi). (I) provide
active transport of therapeutic agents through human and animal GI tissue
(into the blood) for in vivo delivery, particularly for treatment or
prevention of hypertension, diabetes, osteoporosis, haemophilia, anaemia,
cancer, migraine, or angina pectoris. Specifically they are used to

```

```

CC deliver insulin or leuprolide, but many other suitable therapeutic agents
CC are disclosed, including genes or inhibitory nucleic acid, imaging agents
CC and antigens. (I) may also provide targeting to the GI tract. Other uses
CC of (I) are: (i) to determine the level of specified receptors in a sample
CC (in a binding assay); and (ii) to screen for molecules that bind (I).
CC Immunogenic analogues or derivatives of (I) are used to raise antibodies
CC and in immunoassays. The antibodies are used to locate, detect and
CC measure (I), e.g. for imaging, monitoring treatment, tissue analysis
CC etc., also for peptide purification and immobilisation.
SQ Sequence 257 AA:

Query Match 66.8%; Score 1235; DB 20; Length 257;
Best Local Similarity 95.5%; Pred. No. 2, 2e-101;
Matches 231; Conservative 3; Mismatches 6; Indels 2; Gaps 1;

QY 1 MSPIIGYWKIKGLVOPTRLLLEYLEEKYEDEHLYERDEGDKWRNKKFELGEPNLPYYID 60
DB 1 mspilgywkikglvoptrllleyleekyeehlyerdegdkwrnkkfelgfeplpyid 60
QY 61 GDVKLTQSMATIRYIADKHNMIGSCPERAEISMLEGAVIDRYGYSRTAYSKDEFELKY 120
DB 61 gdwkltqsmatiryadi khnmigscperaeismlegavldrygysrtaysdfeclky 120
QY 121 DFLSKLPEMLKMFEDRLCHRTYLNCDHVTHPDFMLYDALDVLYMDPMCLDAPKLVCFK 180
DB 121 df1sklpemlkmedfclchktylngdhvchpdtmlydaldvlymdpmcldapklvcfk 180
QY 181 KRIEALPQIDKYLKSSKYIAMPLOQWQATFGGDPKPSDLVPRGSPNSRVDSGSAEIT 240
DB 181 kriealpqidkylksskylawplqgwqatf9g9dhpksdlvprgsps--lpsvtraaas 238
QY 241 RI 242
DB 239 rv 240

RESULT 4
W74123
W74123 standard; peptide: 282 AA.
W74123:
04-MAY-1999 (first entry)
GST/GI transport receptor binding protein fusion sequence.
Gastro-intestinal transport receptor; binding protein; hsi; HPT1;
D2H; hPEPT1; human; GI tract receptor; sucrose-isomaltase complex;
intestinal peptide-associated transporter; hypertension; diabetes;
osteoporosis; haemophilia; anaemia; cancer; migraine; angina pectoris;
therapeutic agent delivery; therapy; fusion protein.
Homo sapiens.
W09851325-A2.
19-NOV-1998.
15-MAY-1998; 98WO-US10088.
15-MAY-1997; 97US-0046595.
(CYTO-) CYTOGEN CORP.
(ELAN-) ELAN CORP PLC.
Alvarez VL, Belinka BA, Cagney GM, Carter JM, Lambkin IJ,
Omahony DJ, Patterson CA, Singleton J,
WPI: 1999-009568/01.
New proteins that bind specifically to receptors in the

```

PD	30-JUN-1999.	
XX		
PF	21-DEC-1998;	98EP-0310498
XX		
PR	13-OCT-1998;	98US-0103976

PR 30-MAR-1998; 98US-0008629.
PR 21-MAY-1998; 98US-0079857.
PR 01-JUL-1998; 98US-0086321.
PR 05-AUG-1998; 98US-0091385.
XX
PA (ELIL) LILLY & CO ELI.
XX
PI Becker GW, Butler JP, Hale JE, Heath WF, Helman ML;
PI Schoner BE, Varshavsky AD;
XX WPI; 1999-349240/30.
XX
DR
XX New human protein useful for preventing and treating diabetes and
PT the conditions which result from it
PI Claim 1; Page 30-31; 57pp; English.
PS
XX

The specification describes a functional analogue of beta-1ipotrofin
(BUT). The BUT analogues or their fragments may be administered

cc patients to prevent
cc complications while

Sequence 422 AA.
 CC glucosylated. They may also be used to increase sensitivity to insulin
 CC in patients. The present sequence represents a BLT fusion protein.
 XX

Query Match	66.8%	Score 1235	DB 20	Length 422
Best Local Similarity	80.1%	Prod No 4	1101	

	Conserved	Variable	Matches	Indels	Gaps
1	MSPLIGWKIKGLVPTRLLEYLEEKYEHLHYERDEGDKWRNKKFEIGLEFPNIPYAD	60	27	28	4

61 GDVLTOSMAIIRIADKINMLGCCPKRAEISMLGAVLDIRYGSRIASXDFETLKV 120

QY 121 DELEKLEPMKMEEDRLCHKTYLNGDHWTHRPDMLYDALDVYLYMDPMLDAEPKLYCFK 180

[illegible]

```

QY 241 RPLPKGKSLR-----KALKEHGLEDFLOKOO-----YGISSK-- 274
      | | | | | | | | | | | | | | | | | | | | | | | |
241 qrel-tgrrlreaddaddaddaanaaa|ehb|lvaaktrkayruumohfwvgrashah 200

```

QY	275	-YSGFV	279	*
	1	1		
	1	1		
Db	300	ryggim	305	

RESULT	6
MA7001	

AA
AC
XX
W47001:
13-APR-1998 15:00Z

DE
XX
Glutathione-S-transferase and hTERT fusion protein 1.

CC polypeptides of the invention are useful for the diagnosis of Chagas
 CC disease (American Trypanosomiasis), they are capable of detecting
 CC anti-T. cruzi antibodies; or for blood screening. The TCR27 protein
 CC has epitopes to which most T. cruzi infected individuals have
 CC antibodies. The TCR27 polypeptides will not react with serum from
 CC patients with leishmaniasis, schistosomiasis, or autoimmune disease
 CC and are hence less likely to cause false positives in diagnosis.

XX Sequence 564 AA:

Query Match 66.7%; Score 1234.5; DB 16; Length 564;
 Best Local Similarity 89.3%; Pred. No. 6.6e-101;
 Matches 233; Conservative 10; Mismatches 9; Indels 9; Gaps 2;

QY 1 MSPILGYWKIKGLVPTRLLEYLEEKEEHLIERDEGDKRNKKFELGLEPMLPYID 60
 DB 1 MSPILGYWKIKGLVPTRLLEYLEEKEEHLIERDEGDKRNKKFELGLEPMLPYID 60
 QY 61 GDVKLTQSMALIRYIADKINMLGGCPKERAETSMLEGAVALDIRYGSRIASKPEETLV 120
 DB 61 GDVKLTQSMALIRYIADKINMLGGCPKERAETSMLEGAVALDIRYGSRIASKPEETLV 120
 QY 121 DFLSKPEMLKMFEDRLCHKTYLNGDHVTHPDMLYDALDVLYLMDPCLDAFPLKVCFK 180
 DB 121 DFLSKPEMLKMFEDRLCHKTYLNGDHVTHPDMLYDALDVLYLMDPCLDAFPLKVCFK 180
 QY 181 KRIEAIPOIDKYLKSSKTYIAMPLOGWATFGGCHPCKSDLVPRGSPNSRYD----SSG 235
 DB 181 KRIEAIPOIDKYLKSSKTYIAMPLOGWATFGGCHPCKSDLVPRGSPNSRYD----SSG 235
 QY 236 RAEITRIPLKYGKSLRKALKE 256
 DB 241 KMENTKL-----REYVKAKE 257

RESULT 8

R84568
 ID R84568 standard; Protein: 643 AA.

XX R84568;

DT 09-MAY-1996 (first entry)

DE Trypanosoma cruzi TCR27 polypeptide, Ag15.

KW TCR27; Chagas disease; repeat unit; diagnosis; blood screening;
 KW recombinant; fusion protein; glutathione-S-transferase.

OS Trypanosoma cruzi.

XX Key Location/Qualifiers
 FH 329..552
 FT Region /label=repeat_region

FT /note="16 of 69 repeat units of 14 amino acids"

PN W09525797-A1.

XX 28-SEP-1995.

XX 20-MAR-1995; 95WO-US03191.

XX 24-MAR-1994; 94US-0216894.

XX (KIRC/) KIRCHHOFF L V.

XX (OTSU/) OTSU K.

XX Kirchhoff LV, Otsu K;

XX WPI; 1995-344618/44.

XX N-PSDB; T05332.

PT New polypeptide(s) from the TCR27 protein of Trypanosoma cruzi - as

PT immunoassay reagent for specific diagnosis of Chagas disease, also
 PT related nucleic acid and transformed cells

XX Disclosure; Page 40-41; 68pp; English.

CC R84565-R84569 are polypeptides of the TCR27 protein of T. cruzi
 CC The proteins are all fusion products with glutathione-S-transferase
 CC (GST) and some contain a linker sequence. The TCR27 protein comprises
 CC a 95 amino acid (aa) N-terminal region; 69 repeats of a highly
 CC conserved 14 aa sequence, and a 68 aa C-terminal region. This sequence
 CC encodes the GST sequence, the Ag44 polypeptide contg. 16 of the 69
 CC repeat units and also contains the amino and carboxy terminal
 CC peptides of TCR27. The TCR27 polypeptides of the invention are useful
 CC for the diagnosis of Chagas disease (American Trypanosomiasis), they
 CC are capable of detecting anti-T. cruzi antibodies; or for blood
 CC screening. The TCR27 protein has epitopes to which most T. cruzi
 CC infected individuals have antibodies. The TCR27 polypeptides will not
 CC react with serum from patients with leishmaniasis, schistosomiasis,
 CC or autoimmune disease and are hence less likely to cause false
 CC positives in diagnosis.

XX Sequence 643 AA:

Query Match 66.7%; Score 1234.5; DB 16; Length 643;
 Best Local Similarity 89.3%; Pred. No. 7.8e-101;
 Matches 233; Conservative 10; Mismatches 9; Indels 9; Gaps 2;

QY 1 MSPILGYWKIKGLVPTRLLEYLEEKEEHLIERDEGDKRNKKFELGLEPMLPYID 60
 DB 1 MSPILGYWKIKGLVPTRLLEYLEEKEEHLIERDEGDKRNKKFELGLEPMLPYID 60
 QY 61 GDVKLTQSMALIRYIADKINMLGGCPKERAETSMLEGAVALDIRYGSRIASKPEETLV 120
 DB 61 GDVKLTQSMALIRYIADKINMLGGCPKERAETSMLEGAVALDIRYGSRIASKPEETLV 120
 QY 121 DFLSKPEMLKMFEDRLCHKTYLNGDHVTHPDMLYDALDVLYLMDPCLDAFPLKVCFK 180
 DB 121 DFLSKPEMLKMFEDRLCHKTYLNGDHVTHPDMLYDALDVLYLMDPCLDAFPLKVCFK 180
 QY 181 KRIEAIPOIDKYLKSSKTYIAMPLOGWATFGGCHPCKSDLVPRGSPNSRYD----SSG 235
 DB 181 KRIEAIPOIDKYLKSSKTYIAMPLOGWATFGGCHPCKSDLVPRGSPNSRYD----SSG 235
 QY 236 RAEITRIPLKYGKSLRKALKE 256
 DB 241 KMENTKL-----REYVKAKE 257

RESULT 9

W74096
 ID W74096 standard; peptide; 282 AA.

XX W74096;

DT 04-MAY-1999 (first entry)

DE GST/GI transport receptor binding protein fusion sequence.

KW Gastro-intestinal transport receptor; binding protein; hsi; HPT1;
 KW DZH; hPEPT1; human; GI tract receptor; sucrose-isomaltase complex;
 KW intestinal peptide-associated transporter; hypertension; diabetes;
 KW osteoporosis; haemophilia; anaemia; cancer; migraine; angina pectoris;
 KW therapeutic agent delivery; therapy; fusion protein.

XX Homo sapiens.

XX W09851325-A2.

XX 19-NOV-1998.

XX 15-MAY-1998; 98WO-US10088.

PR 15-MAY-1997; 97US-0046595.
 XX (CYTO-) CYTOGEN CORP.
 PA (ELAN-) ELAN CORP PLC.
 XX
 PI Alvarez VL, Belinka BA, Cagney GM, Carter JM, Lambkin LJ;
 PI Omahony DJ, Patterson CA, Singleton J;
 XX WPI: 1999-009568/01.
 DR
 XX
 PT New proteins that bind specifically to receptors in the
 PT gastro-intestinal tract and related nucleic acid - chimaeras and
 PT antibodies, used to deliver therapeutic or diagnostic agents to, or
 PT through, the gastrointestinal tract, e.g. insulin or leuprolide
 XX
 PS Example: Page 204; 294pp; English.

CC This sequence represents a fusion protein between
 CC glutathione-S-transferase and a gastro-intestinal transport protein
 CC binding peptide. The invention relates to purified proteins (I) that bind
 CC specifically to at least one of the gastro-intestinal (GI) tract
 CC receptors human intestinal peptide-associated transporter (HPT1),
 CC hPRT1, D2H and human sucrose-isomaltase complex (hSI). (I) provide
 CC active transport of therapeutic agents through human and animal GI tissue
 CC (into the blood) for in vivo delivery, particularly for treatment or
 CC prevention of hypertension, diabetes, osteoporosis, haemophilia, anaemia,
 CC cancer, migraine, or angina pectoris. Specifically they are used to
 CC deliver insulin or leuprolide, but many other suitable therapeutic agents
 CC are disclosed. Including genes or inhibitory nucleic acid, imaging agents
 CC and antigens. (I) may also provide targeting to the GI tract. Other uses
 CC of (I) are: (i) to determine the level of specified receptors in a sample
 CC (in a binding assay); and (ii) to screen for molecules that bind (i).
 CC Immunogenic analogues or derivatives of (I) are used to raise antibodies
 CC and in immunoassays. The antibodies are used to locate, detect and
 CC measure (i), e.g. for imaging, monitoring treatment, tissue analysis
 CC etc., also for peptide purification and immobilisation.
 XX
 SQ Sequence 282 AA:

Query Match 66.7%; Score 1234; DB 20; Length 282;
 Best Local Similarity 77.8%; Pred. No. 3e-101;
 Matches 245; Conservative 4; Mismatches 20; Indels 46; Gaps 6;
 QY 1 MSPILGYMKIKGLVOTRLLLEYLEEKYEENLYERDEGDMKMKPELLEFPNLPYYID 60
 DB 1 MSPILGYMKIKGLVOTRLLLEYLEEKYEENLYERDEGDMKMKPELLEFPNLPYYID 60
 QY 61 GNVKLTSKAITIRYADKNNMLGSCPKERAETSMLEGAVALDIRGVSRATYSKDFETLV 120
 DB 61 gnvkltskaimlrryladknnmlgscpkeraetlsmlegavldlrrgvsrlayskdfetlv 120
 QY 121 DFLSKPEMLKMFEDRLCKITLNGDHVTHNRPMLYDALDVVLVMPKMLDAFPKLVCKR 180
 DB 121 dflskpemelkmfedrlckitlmgdhvthnrfmlydalvvlympkmldafrpklvckr 180
 QY 181 KRLEIAPIQIDKLLKSSKYTAMPLOGWQATFGGADHPKSDLVPRGSRNVRDSSGRAEIT 240
 DB 181 krlleapldkylksskytamplogwqatfggadhprksdlvprgsrnrsvdssgraet 240
 QY 241 RILPLVKGSLKRLAKLEHGLEDFLOKQOYGISKYSGFVAYTD-CTESGQNLCLCGSNV 299
 DB 241 rllplvkgslkrlaklehgldflokqoygiskysgfvaytd-ctesgqnlclcgsnv 299
 QY 300 CGGKNKCLIGSDGEK 314
 DB 300 cggknkcligsdgek 314
 QY 261 cgnqn-----stgrk 270
 DB 261 cgnqn-----stgrk 270

RESULT 10
 Y52288
 ID Y52288 standard: Protein; 824 AA.
 XX

AC Y52288;
 XX
 DT 09-FEB-2000 (first entry)
 XX
 DE Human HCP/GST fusion protein.
 XX
 KW Haematopoietic cell phosphatase; HCP; cloning; expression;
 KW epithelial cell; growth factor receptor; interleukin-3; IL-3;
 KW erythropoietin; EPO; negative regulator; signal transduction;
 KW modulation; activity; inhibitor; stimulation; cytokine therapy;
 KW antisense therapy; gene therapy; treatment; diagnosis;
 KW glutathione-S-transferase; fusion.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO954450-A1.
 XX
 PD 28-OCT-1999.
 XX
 PF 15-APR-1999; 99WO-US08228.
 XX
 PR 17-APR-1998; 98US-0082072.
 XX
 DR (ORTH) ORTHO-MCNEIL PHARM INC.
 XX
 PI Jolliffe L, Barbone F, Shanahan M, Xu D;
 XX WPI: 2000-013244/01.
 XX N-PSDB: Z32916.
 PT
 PS Assay methods for modulators of hematopoietic cell phosphatase -
 XX
 XX Claim 3; Fig 4B; 57pp; English.
 CC This sequence represents a full-length human haematopoietic
 CC cell phosphatase (HCP)/glutathione-S-transferase (GST) fusion
 CC protein. This was expressed in E. coli and subsequently
 CC purified for use in a variety of activity assays.
 CC HCP is a 68 kd non-membrane associated protein found primarily in
 CC haematopoietic cells, and to some extent in epithelial cells. It
 CC has been shown to associate with several growth factor receptors
 CC following ligand stimulation and acts as a negative regulator
 CC of receptor activation, and hence of haematopoietic signal transduction.
 CC The receptors it regulates include the interleukin-3 (IL-3) and
 CC erythropoietin (EPO) receptors. The invention relates to a novel
 CC method of identifying compounds that modulate HCP activity. This
 CC comprises combining a modulator of HCP protein activity with HCP protein
 CC and HCP protein substrate, and measuring an effect of the modulator on
 CC the activity of the HCP protein. Inhibitors of HCP can be used to
 CC modulate the activity of haematopoietic receptors. HCP has been shown to
 CC be a negative regulator of one or more signal transduction pathways in
 CC haematopoietic cells, and the identification of a HCP inhibitor would
 CC provide a synthetic stimulator to the haematopoietic system could be
 CC used in conjunction with other cytokine therapy, e.g., administration of
 CC EPO. HCP is also associated with several growth factors found in
 CC haematopoietic cells and it is possible that a HCP inhibitor may also
 CC function as a haematopoietic cell potentiator, enhancing the effect of
 CC growth factors, decreasing the necessary dose of the factor. The HCP
 CC polynucleotides are sources of probes and primers, and can be used to
 CC design antisense sequences, and in gene therapy. The protein, or
 CC active fragments thereof, may be used to treat or diagnose disorders
 CC which require the modulation of the HCP activity. Small molecules that
 CC inhibit the action of HCP can also augment the effect of EPO, which
 CC meaning that a lower therapeutic dose of EPO may be utilised.
 XX
 XX Sequence 824 AA:

Query Match 66.7%; Score 1233.5; DB 21; Length 824;
 Best Local Similarity 81.9%; Pred. No. 1.3e-100;
 Matches 245; Conservative 7; Mismatches 24; Indels 23; Gaps 6;

PT antibodies, used to deliver therapeutic or diagnostic agents to, or
 PT through, the gastrointestinal tract, e.g. insulin or leuprolide
 PS Example: Page 226; 294pp; English.

CC This sequence represents a fusion protein between
 CC glutathione-S-transferase and a gastro-intestinal transport protein
 CC binding peptide. The invention relates to purified proteins (I) that bind
 CC specifically to at least one of the gastro-intestinal (GI) tract
 CC receptors human intestinal peptide-associated transporter (HPT1),
 CC hPPT1, D2H and human sucrose-isomaltase complex (HSI). (I) provide
 CC active transport of therapeutic agents through human and animal GI tissue
 CC (into the blood) for in vivo delivery, particularly for treatment or
 CC prevention of hypertension, diabetes, osteoporosis, haemophilia, anaemia,
 CC cancer, migraine, or angina pectoris. Specifically they are used to
 CC deliver insulin or leuprolide, but many other suitable therapeutic agents
 CC are disclosed, including genes or inhibitory nucleic acid, imaging agents
 CC and antigens. (I) may also provide targeting to the GI tract. Other uses
 CC of (I) are: (i) to determine the level of specified receptors in a sample
 CC (in a binding assay); and (ii) to screen for molecules that bind (I).
 CC Immunogenic analogues or derivatives of (I) are used to raise antibodies
 CC and in immunoassays. The antibodies are used to locate, detect and
 CC measure (I), e.g. for imaging, monitoring treatment, tissue analysis
 CC etc., also for peptide purification and immobilisation.

XX Sequence 282 AA;

Query Match 66.6%; Score 1232.5; DB 20; Length 282;
 Best Local Similarity 94.7%; Pred. No. 4.1e-101;
 Matches 233; Conservative 2; Mismatches 6; Indels 5; Gaps 2;

OY 1 MSPILGYWKIKGLVOPTRLLLEYLEEKYEHEHYERDEGDKWRKKFELGLEPPLPYIID 60
 DB 1 MSPILGYWKIKGLVOPTRLLLEYLEEKYEHEHYERDEGDKWRKKFELGLEPPLPYIID 60
 OY 61 GGVKLTQSMATIRYIADKHNMLGCGCKERAETISMEGAVLDIRYGSRIAYSKDFETLVK 120
 DB 61 GGVKLTQSMATIRYIADKHNMLGCGCKERAETISMEGAVLDIRYGSRIAYSKDFETLVK 120
 OY 121 DFLSKLPEMLKMPEDRLCHHTYTLNGDHYHPDMLVDALDVVLYMPMCLDAPKLVCFK 180
 DB 121 DFLSKLPEMLKMPEDRLCHHTYTLNGDHYHPDMLVDALDVVLYMPMCLDAPKLVCFK 180
 OY 181 KRLEAIPOIDKYLKSSKYIAMPLOGQATFGGDHPKSDLVPRGSPNSRVDSSGRAEIT 240
 DB 181 KRLEAIPOIDKYLKSSKYIAMPLOGQATFGGDHPKSDLVPRGSPNSRVDSSGRAEIT 240
 OY 241 RIPLVYK 246
 DB 239 r---yk 241

RESULT 13
 W74113 standard; peptide: 248 AA.

AC W74113:
 XX 04-MAY-1999 (first entry)

DE GST/GI transport receptor binding protein fusion sequence.
 XX Gastro-intestinal transport receptor; binding protein; hsi: HPT1;
 XX D2H, hPPT1, human, GI tract receptor; sucrose-isomaltase complex;
 KW intestinal peptide-associated transporter; hypertension; diabetes;
 KW osteoporosis; haemophilia; anaemia; cancer; migraine; angina pectoris;
 XX therapeutic agent delivery; therapy; fusion protein.

OS Homo sapiens.
 XX
 PN W09051325-A2.
 XX

PD 19-NOV-1998.
 XX
 PF 15-MAY-1998: 98WO-US10088.
 XX
 PR 15-MAY-1997: 97US-0046595.

PA (CYTO-) CYTOGEN CORP.
 PA (ELAN-) ELAN CORP PLC.
 PI Alvarez VL, Belinka BA, Cagney GW, Carter JM, Lambkin IJ;
 PI Omahony DJ, Patterson CA, Singleton J;
 DR WPI: 1999-009568/01.

PT New proteins that bind specifically to receptors in the
 PT gastro-intestinal tract and related nucleic acid - chimaeras and
 PT antibodies, used to deliver therapeutic or diagnostic agents to, or
 PT through, the gastrointestinal tract, e.g. insulin or leuprolide
 PS Example: Page 216; 294pp; English.

CC This sequence represents a fusion protein between
 CC glutathione-S-transferase and a gastro-intestinal transport protein
 CC binding peptide. The invention relates to purified proteins (I) that bind
 CC specifically to at least one of the gastro-intestinal (GI) tract
 CC receptors human intestinal peptide-associated transporter (HPT1),
 CC hPPT1, D2H and human sucrose-isomaltase complex (HSI). (I) provide
 CC active transport of therapeutic agents through human and animal GI tissue
 CC (into the blood) for in vivo delivery, particularly for treatment or
 CC prevention of hypertension, diabetes, osteoporosis, haemophilia, anaemia,
 CC cancer, migraine, or angina pectoris. Specifically they are used to
 CC deliver insulin or leuprolide, but many other suitable therapeutic agents
 CC are disclosed, including genes or inhibitory nucleic acid, imaging agents
 CC and antigens. (I) may also provide targeting to the GI tract. Other uses
 CC of (I) are: (i) to determine the level of specified receptors in a sample
 CC (in a binding assay); and (ii) to screen for molecules that bind (I).
 CC Immunogenic analogues or derivatives of (I) are used to raise antibodies
 CC and in immunoassays. The antibodies are used to locate, detect and
 CC measure (I), e.g. for imaging, monitoring treatment, tissue analysis
 CC etc., also for peptide purification and immobilisation.

XX Sequence 248 AA;

Query Match 66.6%; Score 1232; DB 20; Length 248;
 Best Local Similarity 95.9%; Pred. No. 3.9e-101;
 Matches 231; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

OY 1 MSPILGYWKIKGLVOPTRLLLEYLEEKYEHEHYERDEGDKWRKKFELGLEPPLPYIID 60
 DB 1 MSPILGYWKIKGLVOPTRLLLEYLEEKYEHEHYERDEGDKWRKKFELGLEPPLPYIID 60
 OY 61 GGVKLTQSMATIRYIADKHNMLGCGCKERAETISMEGAVLDIRYGSRIAYSKDFETLVK 120
 DB 61 GGVKLTQSMATIRYIADKHNMLGCGCKERAETISMEGAVLDIRYGSRIAYSKDFETLVK 120
 OY 121 DFLSKLPEMLKMPEDRLCHHTYTLNGDHYHPDMLVDALDVVLYMPMCLDAPKLVCFK 180
 DB 121 DFLSKLPEMLKMPEDRLCHHTYTLNGDHYHPDMLVDALDVVLYMPMCLDAPKLVCFK 180
 OY 181 KRLEAIPOIDKYLKSSKYIAMPLOGQATFGGDHPKSDLVPRGSPNSRVDSSGRAEIT 240
 DB 181 KRLEAIPOIDKYLKSSKYIAMPLOGQATFGGDHPKSDLVPRGSPNSRVDSSGRAEIT 240
 OY 241 R 241
 DB 239 r 239

RESULT 14
 W74122 standard; peptide: 248 AA.
 ID W74122
 XX

AC W74122;
 XX
 DT 04-MAY-1999 (first entry)
 XX
 DE GST/GI transport receptor binding protein fusion sequence.
 XX
 KM Gastro-intestinal transport receptor; binding protein; hSI; HPT1;
 KM D2H; hPEPT1; human; GI tract receptor; sucrose-isomaltase complex;
 KM intestinal peptide-associated transporter; hypertension; diabetes;
 KM osteoporosis; haemophilia; anaemia; cancer; migraine; angina pectoris;
 KM therapeutic agent delivery; therapy; fusion protein.
 XX
 OS Homo sapiens.
 XX
 PN WO9851325-A2.
 XX
 PD 19-NOV-1998.
 XX
 PF 15-MAY-1998; 98WO-US10088.
 XX
 PR 15-MAY-1997; 97US-0046595.
 XX
 PA (CYTO-) CYTOGEN CORP.
 PA (ELAN-) ELAN CORP PLC.
 XX
 PI Alvarez VL, Belinka BA, Cagney GM, Carter JM, Lambkin IJ;
 PI Omahony DJ, Patterson CA, Singleton J;
 XX
 DR WPI: 1999-009568/01.
 XX
 PT New proteins that bind specifically to receptors in the
 PT gastro-intestinal tract and related nucleic acid - chimeras and
 PT antibodies, used to deliver therapeutic or diagnostic agents to, or
 PT through, the gastrointestinal tract, e.g. insulin or leuprolide
 XX
 PS Example: Page 222-223; 294pp; English.
 XX
 CC This sequence represents a fusion protein between
 CC glutathione-S-transferase and a gastro-intestinal transport protein
 CC binding peptide. The invention relates to purified proteins (I) that bind
 CC specifically to at least one of the gastro-intestinal (GI) tract
 CC receptors human intestinal peptide-associated transporter (hPT1),
 CC hPEPT1, D2H and human sucrose-isomaltase complex (hSI). (I) provide
 CC active transport of therapeutic agents through human and animal GI tissue
 CC (into the blood) for in vivo delivery, particularly for treatment or
 CC prevention of hypertension, diabetes, osteoporosis, haemophilia, anaemia,
 CC cancer, migraine, or angina pectoris. Specifically they are used to
 CC deliver insulin or leuprolide, but many other suitable therapeutic agents
 CC are disclosed. Including genes or inhibitory nucleic acid, imaging agents
 CC and antigens. (I) may also provide targeting to the GI tract. Other uses
 CC of (I) are: (i) to determine the level of specified receptors in a sample
 CC (in a binding assay); and (ii) to screen for molecules that bind (I).
 CC Immunogenic analogues or derivatives of (I) are used to raise antibodies
 CC and in immunoassays. The antibodies are used to locate, detect and
 CC measure (I), e.g. for imaging, monitoring treatment, tissue analysis
 CC etc., also for peptide purification and immobilisation.
 XX
 SQ Sequence 248 AA;
 Query Match 66.6%; Score 1232; DB 20; Length 248;
 Best Local Similarity 95.9%; Pred. No. 3,9e-10;
 Matches 231; Conservative 2; Mismatches 6; Indels 2; Gaps 1;
 QY 1 MSPIIGWYKIKLVNPTRLLEKYEHEHYERDEGDKMNKKFEELGEPNIPYVD 60
 DB 1 MSPILGYWIKIVJPTLILEYEEKYEELHYERDEGDKMNKKFEELGEPNIPYVD 60
 QY 61 GGVKLTQSMAIRYADKNNMLGCPKERAETSMLEGAVALDRYGVSRAYSKEDEETLKV 120
 DB 61 ggvkltqsmairyladknnmlgcpkeraetasmlegavaldrryvsrslayskedettlkv 120
 QY 121 DFLSKLPENLKMFEEDRLCHKTYLNGDHTHPDMLYDALDVVLYMDPKCLAPKLVCFK 180

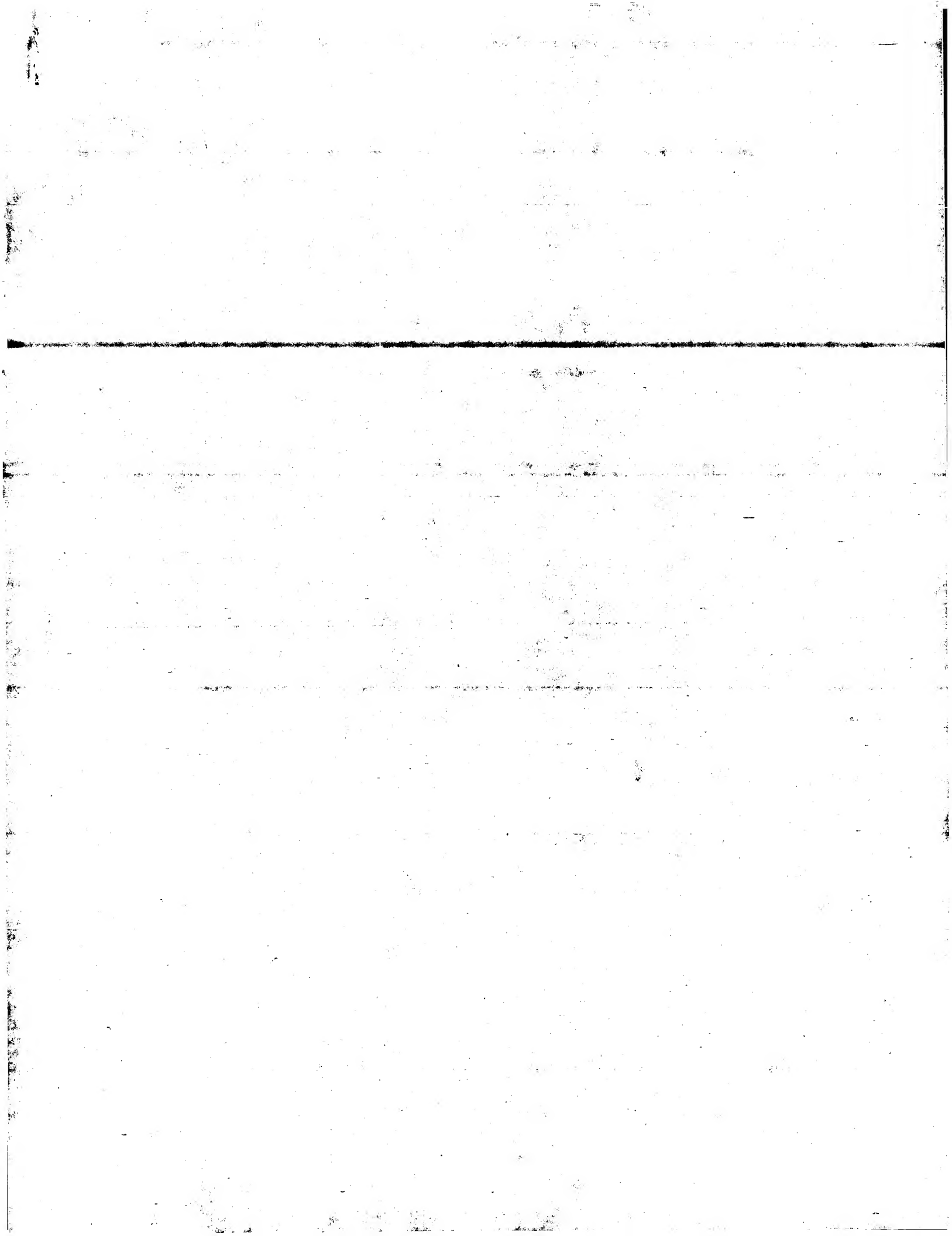
DB 121 dflsklpemlkmfedrlchkyngdhvthpdmlydaldvlymdpncldafklycfk 180
 QY 181 KRFAIPQIDKYLKSSKTIAMPLOGWATFEGGCHPPEKSDLVPPGSRPSRDSSGRAEIT. 240
 DB 181 krfaipqidkylkssktyamplogwatfeggchpppekstdlvppgsrpsrdssgraeit. 238
 QY 241 R 241
 DB 239 r 239
 RESULT 15
 W74117
 ID W74117 standard; peptide; 258 AA.
 XX
 AC W74117;
 XX
 DT 04-MAY-1999 (first entry)
 XX
 DE GST/GI transport receptor binding protein fusion sequence.
 XX
 KM Gastro-intestinal transport receptor; binding protein; hSI; HPT1;
 KM D2H; hPEPT1; human; GI tract receptor; sucrose-isomaltase complex;
 KM intestinal peptide-associated transporter; hypertension; diabetes;
 KM osteoporosis; haemophilia; anaemia; cancer; migraine; angina pectoris;
 KM therapeutic agent delivery; therapy; fusion protein.
 XX
 OS Homo sapiens.
 XX
 PN WO9851325-A2.
 XX
 PD 19-NOV-1998.
 XX
 PF 15-MAY-1998; 98WO-US10088.
 XX
 PR 15-MAY-1997; 97US-0046595.
 XX
 PA (CYTO-) CYTOGEN CORP.
 PA (ELAN-) ELAN CORP PLC.
 XX
 PI Alvarez VL, Belinka BA, Cagney GM, Carter JM, Lambkin IJ;
 PI Omahony DJ, Patterson CA, Singleton J;
 XX
 DR WPI: 1999-009568/01.
 XX
 PT New proteins that bind specifically to receptors in the
 PT gastro-intestinal tract and related nucleic acid - chimeras and
 PT antibodies, used to deliver therapeutic or diagnostic agents to, or
 PT through, the gastrointestinal tract, e.g. insulin or leuprolide
 XX
 PS Example: Page 219; 294pp; English.
 XX
 CC This sequence represents a fusion protein between
 CC glutathione-S-transferase and a gastro-intestinal transport protein
 CC binding peptide. The invention relates to purified proteins (I) that bind
 CC specifically to at least one of the gastro-intestinal (GI) tract
 CC receptors human intestinal peptide-associated transporter (hPT1),
 CC hPEPT1, D2H and human sucrose-isomaltase complex (hSI). (I) provide
 CC active transport of therapeutic agents through human and animal GI tissue
 CC (into the blood) for in vivo delivery, particularly for treatment or
 CC prevention of hypertension, diabetes, osteoporosis, haemophilia, anaemia,
 CC cancer, migraine, or angina pectoris. Specifically they are used to
 CC deliver insulin or leuprolide, but many other suitable therapeutic agents
 CC are disclosed. Including genes or inhibitory nucleic acid, imaging agents
 CC and antigens. (I) may also provide targeting to the GI tract. Other uses
 CC of (I) are: (i) to determine the level of specified receptors in a sample
 CC (in a binding assay); and (ii) to screen for molecules that bind (I).
 CC Immunogenic analogues or derivatives of (I) are used to raise antibodies
 CC and in immunoassays. The antibodies are used to locate, detect and
 CC measure (I), e.g. for imaging, monitoring treatment, tissue analysis
 CC etc., also for peptide purification and immobilisation.
 XX

50 Sequence 258 AA;

Query Match	66.6%;	Score 1232;	DB 20;	Length 258;
Best Local Similarity	95.1%;	Pred. No. 4.1e-101;		
Matches 231; Conservative	2;	Mismatches 8;	Indels 2;	Gaps 1;

QY	1	MSPIIGGWKIGGLVQPRRLLEVEEKEEHLVERDEGDKWRNKKFELGLEPNNLYID	60
Db	1	msp1llygwklglvqprl1leyeekeehlyerdegdwrrnkkfelglefnpnlyid	60
QY	61	GDVKTQSMALIRYIADKHNNLGGCPKERAEMISMLGCAVLDIRGVSRVAYSKDPEFLKV	120
Db	61	gdvktltsmaliriyiadkhnmlggcpkeraeismlegavidirgvsrvtlayskdpeflkv	120
QY	121	DFLSKLEPMKMFEDRLCHHTYTLNGDHVTHPPDMLVDALDVLVYMPMCDARPKVCRK	180
Db	121	df1sk1pema1kmedr1chtktytlngdhvthppdmlvda1dv1vympmc1darpk1vcrk	180
QY	181	KRIEAIPOIDKYLKSSKYIAMPLOGWATFGGDHPRSDLVBERGSPNRSVDSGSAEIT	240
Db	181	krieaipo1dkylksskyiamplogwatfggdhprsd1vbergs1pnsrsvds1gsae1t	240
QY	241	RIP	243
Db	239	stp	241

Search completed: March 18, 2001, 05:32:01
Job time: 2079 sec



Db 61 GDVKLTQSMALLIRYADKHNMLGCGPKERAETSMLEGAVIDIRYGVSRIVASKDEFTLKV 120
QY 121 DFLSLPMLKMFEDRLCHKTYLNDHTHPDMYDALDVLYMDPCLDAFPLVCEK 180
Db 121 DFLSLPMLKMFEDRLCHKTYLNDHTHPDMYDALDVLYMDPCLDAFPLVCEK 180
QY 181 KRIEAIPOIDKYLYKSSKIYAMPLOGQWATFGGSDHPKPSDLPVPGSPRSRYDSSG 235
Db 181 KRIEAIPOIDKYLYKSSKIYAMPLOGQWATFGGSDHPKPSDLPVPGSPRSRYDSSG 240
QY 236 RAETRIPLVYKGSILRKALKE 256
Db 241 KKEMTKL---REKYKKAKE 257

RESULT 2
US-08-974-549A-602
Sequence 602, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cecch, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-00261005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 602:
SEQUENCE CHARACTERISTICS:
LENGTH: 538 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..538
OTHER INFORMATION: glutathione-S-transferase sequence,
OTHER INFORMATION: thrombin cleavage sequence, recognition
OTHER INFORMATION: sequence for heart muscle protein
OTHER INFORMATION: kinase, residues introduced by cloning
OTHER INFORMATION: and hTERT protein fragment"
US-08-974-549A-602

Query Match 66.7%; Score 1234.5; DB 3; Length 538;
Best Local Similarity 85.1%; Pred. No. 1,6e-105;
Matches 240; Conservative 6; Mismatches 13; Indels 23; Gaps 4;

QY 1 MSPILGYWKIKIGLVOPTRLLLEYLEKEEHLYEDEDECKRNKKFELGFPNLPYID 60
Db 1 MSPILGYWKIKIGLVOPTRLLLEYLEKEEHLYEDEDECKRNKKFELGFPNLPYID 60
QY 61 GDVKLTQSMALLIRYADKHNMLGCGPKERAETSMLEGAVIDIRYGVSRIVASKDEFTLKV 120
Db 61 GDVKLTQSMALLIRYADKHNMLGCGPKERAETSMLEGAVIDIRYGVSRIVASKDEFTLKV 120
QY 121 DFLSLPMLKMFEDRLCHKTYLNDHTHPDMYDALDVLYMDPCLDAFPLVCEK 180
Db 121 DFLSLPMLKMFEDRLCHKTYLNDHTHPDMYDALDVLYMDPCLDAFPLVCEK 180
QY 181 KRIEAIPOIDKYLYKSSKIYAMPLOGQWATFGGSDHPKPSDLPVPGSPRSRYDSSG 235
Db 181 KRIEAIPOIDKYLYKSSKIYAMPLOGQWATFGGSDHPKPSDLPVPGSPRSRYDSSG 240
QY 241 RIP-----LYKG---KSLRKALKEHGL-----EDFL 264
Db 236 KIPQGSILSTLCSLCYGDMEKIFAGIRDGLLRVDDFL 277

RESULT 3
US-08-216-894-2
Sequence 2, Application US/08216894
Patent No. 5876734
GENERAL INFORMATION:
APPLICANT: Kirchhoff, Louis V.
APPLICANT: Otsu, Keiko
TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/216,894
FILING DATE: 24-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 85326/102/DRLO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 564 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-216-894-2

Query Match 66.7%; Score 1234.5; DB 2; Length 564;
Best Local Similarity 89.3%; Pred. No. 1.7e-105;
Matches 233; Conservative 10; Mismatches 9; Indels 9; Gaps 2;

QY 1 MSPILGYWKIKGLVOPTRILLEYLEEKYEHLIERDEGDKWRNKKFELGLEFPNLPYYID 60
DB 1 MSPILGYWKIKGLVOPTRILLEYLEEKYEHLIERDEGDKWRNKKFELGLEFPNLPYYID 60
QY 61 GGVKLTQSMALIRYIADKNNMLGGCCKERAETSMLEGAVALDIRYGSRTAYSDFETLKY 120
DB 61 GGVKLTQSMALIRYIADKNNMLGGCCKERAETSMLEGAVALDIRYGSRTAYSDFETLKY 120
QY 121 DFLSKLPEMLKMFEDRLCKHTYLNAGDHVTHPPDMLDALDVLYLMPKCLDAFPKLYCKR 180
DB 121 DFLSKLPEMLKMFEDRLCKHTYLNAGDHVTHPPDMLDALDVLYLMPKCLDAFPKLYCKR 180
QY 181 KRIEALPQIDKYLKSSKYIAMPLOGQATFGGDHPKSDLYVRGSPNSRVD-----SSG 235
DB 181 KRIEALPQIDKYLKSSKYIAMPLOGQATFGGDHPKSDLYVRGSPNSRVD-----SSG 235
QY 236 RAEITRIPLYKSKSLKALKE 256
DB 241 KKEMTKL---REKVKRAEKE 257

RESULT 4

US-08-216-894-8
Sequence 8, Application US/08216894
Patent No. 5876734
GENERAL INFORMATION:
APPLICANT: Kirchoff, Louis V.
APPLICANT: Otsu, Keiko
TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION
TITLE OF INVENTION: WITH TRYPAPOSOMA CRUZI
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/216,894
FILING DATE: 24-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 85326/102/DRLO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 643 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-216-894-8

Query Match 66.7%; Score 1234.5; DB 2; Length 643;
Best Local Similarity 89.3%; Pred. No. 2e-105;
Matches 233; Conservative 10; Mismatches 9; Indels 9; Gaps 2;

QY 1 MSPILGYWKIKGLVOPTRILLEYLEEKYEHLIERDEGDKWRNKKFELGLEFPNLPYYID 60
DB 1 MSPILGYWKIKGLVOPTRILLEYLEEKYEHLIERDEGDKWRNKKFELGLEFPNLPYYID 60
QY 61 GGVKLTQSMALIRYIADKNNMLGGCCKERAETSMLEGAVALDIRYGSRTAYSDFETLKY 120
DB 61 GGVKLTQSMALIRYIADKNNMLGGCCKERAETSMLEGAVALDIRYGSRTAYSDFETLKY 120
QY 121 DFLSKLPEMLKMFEDRLCKHTYLNAGDHVTHPPDMLDALDVLYLMPKCLDAFPKLYCKR 180
DB 121 DFLSKLPEMLKMFEDRLCKHTYLNAGDHVTHPPDMLDALDVLYLMPKCLDAFPKLYCKR 180
QY 181 KRIEALPQIDKYLKSSKYIAMPLOGQATFGGDHPKSDLYVRGSPNSRVD-----SSG 235
DB 181 KRIEALPQIDKYLKSSKYIAMPLOGQATFGGDHPKSDLYVRGSPNSRVD-----SSG 235
QY 236 RAEITRIPLYKSKSLKALKE 256
DB 241 KKEMTKL---REKVKRAEKE 257

RESULT 5

US-08-974-549A-604
Sequence 604, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Langer, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 604:
SEQUENCE CHARACTERISTICS:
LENGTH: 515 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..515
OTHER INFORMATION: /note= "Fusion protein composed of
OTHER INFORMATION: glutathione-S-transferase sequence,
OTHER INFORMATION: thrombin cleavage sequence, recognition
OTHER INFORMATION: sequence for heart muscle protein
OTHER INFORMATION: kinase, residues introduced by cloning
OTHER INFORMATION: and hprt protein fragment"
US-08-974-549A-604

Query Match 66.6%; Score 1232.5; DB 3; Length 515;
Best Local Similarity 92.4%; Pred. No. 2.3e-105;
Matches 232; Conservative 6; Mismatches 8; Indels 5; Gaps 1;

QY 1 MSPILGWIKIGVOPTRILLEYLEERKYEHELYERDEGDKWRNKKPELGEPNLPYYID 60
DB 1 MSPILGWIKIGVOPTRILLEYLEERKYEHELYERDEGDKWRNKKPELGEPNLPYYID 60
QY 61 GDVKLTQSMATIRYIADKNHMLGGCKREAEISMLEGAVLIDITGYSRVIAVSDEFTLVY 120
DB 61 GDVKLTQSMATIRYIADKNHMLGGCKREAEISMLEGAVLIDITGYSRVIAVSDEFTLVY 120
QY 121 DFLSKLPEMKMFEDRCHKTYINGDHVHPDPMLYDALDVLYLMDPMCLDAPPKLVCFK 180
DB 121 DFLSKLPEMKMFEDRCHKTYINGDHVHPDPMLYDALDVLYLMDPMCLDAPPKLVCFK 180
QY 181 KRTEAIPQIDKYLKSSKYIAMPLOGWQATEGGDHPKSDLVPRGSPNSRVSSGRAEIT 240
DB 181 KRTEAIPQIDKYLKSSKYIAMPLOGWQATEGGDHPKSDLVPRGSPNSRVSSGRAEIT 240
QY 241 RPLVYKGSJR 251

Db 236 KMSYVVELLR 246
RESULT 6
US-08-974-549A-606
Sequence 606, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cecchi, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morlin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 606:
SEQUENCE CHARACTERISTICS:

LENGTH: 517 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..517
OTHER INFORMATION: /note="fusion protein composed of
OTHER INFORMATION: glutathione-S-transferase sequence,
OTHER INFORMATION: thrombin cleavage sequence, recognition
OTHER INFORMATION: sequence for heart muscle protein
OTHER INFORMATION: kinase residues introduced by cloning
OTHER INFORMATION: and hprt protein fragment
US-08-974-549A-606

Query Match 66.5%; Score 1229.5; DB 3; Length 517;
Best Local Similarity 94.7%; Pred. No. 4.4e-105;
Matches 230; Conservative 4; Mismatches 4; Indels 5; Gaps 1;

QY 1 MSPILGYWKIKGLVQPTRLLEYLEEKEYEHELYERDEGDKMRKKFELGLEFPNLPYYID 60
DB 1 MSPILGYWKIKGLVQPTRLLEYLEEKEYEHELYERDEGDKMRKKFELGLEFPNLPYYID 60
QY 61 GDVKLTQSMALIRYIADKHNMLGGCPKERAELISMLEGAVLDIRYGVSRAYSXDFETLVY 120
DB 61 GDVKLTQSMALIRYIADKHNMLGGCPKERAELISMLEGAVLDIRYGVSRAYSXDFETLVY 120
QY 121 DFLSKLPEMLKMFEDRLCKHTYINGDHVTHPPDMLYDALDVLVYMPMCLDAPFKLVCKR 180
DB 121 DFLSKLPEMLKMFEDRLCKHTYINGDHVTHPPDMLYDALDVLVYMPMCLDAPFKLVCKR 180
QY 181 KRLEAIPOIDKYLKSSKYIAMPLOGQATFGGSDHPKSDLVPRGSPNSRVDS--SRAEIT 240
DB 181 KRLEAIPOIDKYLKSSKYIAMPLOGQATFGGSDHPKSDLVPRGSPNSRVDS--SRAEIT 240
QY 241 RIP 243
DB 236 KMP 238

RESULT 7
US-08-974-549A-603
Sequence 603, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-00261005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 603:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..530
OTHER INFORMATION: /note="fusion protein composed of
OTHER INFORMATION: glutathione-S-transferase sequence,
OTHER INFORMATION: thrombin cleavage sequence, recognition
OTHER INFORMATION: sequence for heart muscle protein
OTHER INFORMATION: kinase residues introduced by cloning,
OTHER INFORMATION: eight consecutive His residues and hprt
OTHER INFORMATION: protein fragment"
US-08-974-549A-603

Query Match 66.4%; Score 1229; DB 3; Length 530;
Best Local Similarity 90.3%; Pred. No. 5e-105;
Matches 232; Conservative 6; Mismatches 13; Indels 6; Gaps 1;

QY 1 MSPILGYWKIKGLVQPTRLLEYLEEKEYEHELYERDEGDKMRKKFELGLEFPNLPYYID 60
DB 1 MSPILGYWKIKGLVQPTRLLEYLEEKEYEHELYERDEGDKMRKKFELGLEFPNLPYYID 60
QY 61 GDVKLTQSMALIRYIADKHNMLGGCPKERAELISMLEGAVLDIRYGVSRAYSXDFETLVY 120
DB 61 GDVKLTQSMALIRYIADKHNMLGGCPKERAELISMLEGAVLDIRYGVSRAYSXDFETLVY 120
QY 121 DFLSKLPEMLKMFEDRLCKHTYINGDHVTHPPDMLYDALDVLVYMPMCLDAPFKLVCKR 180
DB 121 DFLSKLPEMLKMFEDRLCKHTYINGDHVTHPPDMLYDALDVLVYMPMCLDAPFKLVCKR 180
QY 181 KRLEAIPOIDKYLKSSKYIAMPLOGQATFGGSDHPKSDLVPRGSPNSRVDS-----S 234

Db 181 KRIEAIPOIDKYLKSKYIAMPLOGWQATFGGDHPKPSDLVPRGSPRASYGVHHHHH 240
QY 235 GRAETTRIPILKXGSLR 251
Db 241 HHSVTKMSYVVELLR 257

RESULT 8
PCT-US94-09700-6

Sequence 6, Application PC/TUS9409700
GENERAL INFORMATION:
APPLICANT: BAYLOR COLLEGE OF MEDICINE
APPLICANT: SMITH, JAMES R.
TITLE OF INVENTION: SENESCENT CELL DERIVED INHIBITORS OF
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOWREY & SIMON
STREET: 1299 PENNSYLVANIA AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09700
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/808,523
FILING DATE: 16-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/970,462
FILING DATE: 02-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/113,372
FILING DATE: 30-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/153,564
FILING DATE: 17-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/203,535
FILING DATE: 25-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/229,420
FILING DATE: 15-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/274,535
FILING DATE: 13-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: AUERBACH, JEFFREY I.
REGISTRATION NUMBER: 32,680
REFERENCE/DOCKET NUMBER: 225-102-CIP7-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 383-7451
TELEFAX: (202) 383-6610
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 232 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Schistosoma japonicum
IMMEDIATE SOURCE:
CLONE: GST
PCT-US94-09700-6

Query Match 66.3%; Score 1227; DB 4; Length 232;
Best Local Similarity 98.3%; Pred. No. 2,6e-105;
Matches 228; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MSPILGYWKIKGLVQPTRLLEYLEEYEEHLVRRDEGDKWRNKKFELGLEPMLPYID 60
Db 1 MSPILGYWKIKGLVQPTRLLEYLEEYEEHLVRRDEGDKWRNKKFELGLEPMLPYID 60
QY 61 GDVKLTOSMAIRIYIAKHNMLGCGPKERAISMLEGAVIDIRGVSRIVASKDFETLV 120
Db 61 GDVKLTOSMAIRIYIAKHNMLGCGPKERAISMLEGAVIDIRGVSRIVASKDFETLV 120
QY 121 DFLSKPEMLKMFEDRLCKTYLNGDHYTHDPFMYLDALDVLVMDPCLDAPFKVCFK 180
Db 121 DFLSKPEMLKMFEDRLCKTYLNGDHYTHDPFMYLDALDVLVMDPCLDAPFKVCFK 180
QY 181 KRIEAIPOIDKYLKSKYIAMPLOGWQATFGGDHPKPSDLVPRGSPNSRVD 232
Db 181 KRIEAIPOIDKYLKSKYIAMPLOGWQATFGGDHPKPSDLVPRGSPNSRVD 232

RESULT 9
PCT-US94-09700-11

Sequence 11, Application PC/TUS9409700
GENERAL INFORMATION:
APPLICANT: BAYLOR COLLEGE OF MEDICINE
APPLICANT: SMITH, JAMES R.
TITLE OF INVENTION: SENESCENT CELL DERIVED INHIBITORS OF
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOWREY & SIMON
STREET: 1299 PENNSYLVANIA AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09700
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/808,523
FILING DATE: 16-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/970,462
FILING DATE: 02-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/113,372
FILING DATE: 30-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/153,564
FILING DATE: 17-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/203,535
FILING DATE: 25-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/229,420
FILING DATE: 15-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/274,535
FILING DATE: 13-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: AUERBACH, JEFFREY I.
REGISTRATION NUMBER: 32,680
REFERENCE/DOCKET NUMBER: 225-102-CIP7-PCT

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 383-7451
TELEFAX: (202) 383-6610
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 397 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
HYPOTHETICAL: NO
IMMEDIATE SOURCE:
CLONE: GST-SDI-1 fusion protein
PCT-US94-09700-11

Query Match 66.3%; Score 1226; DB 4; Length 397;
Best Local Similarity 100.0%; Pred. No. 6,5e-105;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSPILGYWKIKGLVPTRLLEYLEEKYEHLHYERDEGDKWRNKKFELGLEFPNLPYYID 60
Db 1 MSPILGYWKIKGLVPTRLLEYLEEKYEHLHYERDEGDKWRNKKFELGLEFPNLPYYID 60
Qy 61 GVKLTQSMATIRYIADKHNMLGGCPKERAISMLEGAVLDIRYGSRIAYSKDFETLKY 120
Db 61 GVKLTQSMATIRYIADKHNMLGGCPKERAISMLEGAVLDIRYGSRIAYSKDFETLKY 120
Qy 121 DFLSKLPEMLKMFEDRLCHRTYLNCDHVTHPDMLYDALDVLYMPMCLDAFPKLVCFK 180
Db 121 DFLSKLPEMLKMFEDRLCHRTYLNCDHVTHPDMLYDALDVLYMPMCLDAFPKLVCFK 180
Qy 181 KRIEAIPOIDKYKSSKYIAMPLOGQWATFGGDDHPKSDLVPRGSP 227
Db 181 KRIEAIPOIDKYKSSKYIAMPLOGQWATFGGDDHPKSDLVPRGSP 227

RESULT 10
US-08-974-549A-605
Sequence 605, Application US/08974549A
Patent No. 6166178

GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-00261005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 605:
SEQUENCE CHARACTERISTICS:
LENGTH: 514 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: Protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..515
OTHER INFORMATION: /note="fusion protein composed of
glutathione-S-transferase sequence,
thrombin cleavage sequence, recognition
OTHER INFORMATION: sequence for heart muscle protein
OTHER INFORMATION: kinase, residues introduced by cloning
OTHER INFORMATION: and hprt protein fragment"

US-08-974-549A-605

Query Match 66.3%; Score 1226; DB 3; Length 514;
Best Local Similarity 96.6%; Pred. No. 9,1e-105;
Matches 229; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MSPILGYWKIKGLVPTRLLEYLEEKYEHLHYERDEGDKWRNKKFELGLEFPNLPYYID 60
Db 1 MSPILGYWKIKGLVPTRLLEYLEEKYEHLHYERDEGDKWRNKKFELGLEFPNLPYYID 60
Qy 61 GVKLTQSMATIRYIADKHNMLGGCPKERAISMLEGAVLDIRYGSRIAYSKDFETLKY 120
Db 61 GVKLTQSMATIRYIADKHNMLGGCPKERAISMLEGAVLDIRYGSRIAYSKDFETLKY 120
Qy 121 DFLSKLPEMLKMFEDRLCHRTYLNCDHVTHPDMLYDALDVLYMPMCLDAFPKLVCFK 180
Db 121 DFLSKLPEMLKMFEDRLCHRTYLNCDHVTHPDMLYDALDVLYMPMCLDAFPKLVCFK 180
Qy 181 KRIEAIPOIDKYKSSKYIAMPLOGQWATFGGDDHPKSDLVPRGSPNSVSSGRA 237
Db 181 KRIEAIPOIDKYKSSKYIAMPLOGQWATFGGDDHPKSDLVPRGSPNSVSSGRA 237

RESULT 11
US-08-864-224-11
Sequence 11, Application US/08864224

Patent No 5851808
GENERAL INFORMATION:
APPLICANT: Elledge, Stephen J.
APPLICANT: Liu, Qinghua
TITLE OF INVENTION: Rapid Subcloning Using Site-Specific
TITLE OF INVENTION: Recombination
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
City: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent'n Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/864,224
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: BCM-02681
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 579 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-864-224-11

ADDRESSEE: Pollock, Vande Sande & Priddy
STREET: 1990 M Street, NW, Suite 800
City: Washington
STATE: DC
COUNTRY: USA
Zip: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/395,507
FILING DATE: 27-FEB-1995
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP Hei-6-54672
FILING DATE: 28-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Amerinick, Burton A.
REGISTRATION NUMBER: 24,952
REFERENCE/DOCKET NUMBER: 1581/007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 331-7111
TELEFAX: (202) 293-6229
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
STRANDNESS:
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: rabbit
FEATURE:
NAME/KEY: Protein
LOCATION: 1..352
OTHER INFORMATION: /note- "G15 Antigen"
US-08-395-507-1

RESULT 12
 US-08-395-507-1
 : Sequence 1, Application US/08395507
 : Patent No. 5578456
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Fujimura, Katsuya
 :
 : APPLICANT: Ueno, Eichi
 :
 : APPLICANT: Fujii, No. 5578456yuki
 :
 : APPLICANT: Okada, Masahisa
 :
 : TITLE OF INVENTION: Anti-treponema pallidum Antibody
 :
 : TITLE OF INVENTION: Immunoassay
 :
 : NUMBER OF SEQUENCES: 2
 :
 : CORRESPONDENCE ADDRESSES:

[illegible]

```

APPLICANT: Guemene, Daniel
APPLICANT: Zadowny, David
APPLICANT: Karatzas, Costas
TITLE OF INVENTION: USE OF PROLACTIN FOR PREVENTING AND/OR
TITLE OF INVENTION: TREATING BIRD BROODINESS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESS: WEISER & ASSOCIATES
STREET: 230 South Fifteenth Street, Suite 500
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/37,248
FILING DATE: 28-APR-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR95/00576
FILING DATE: 03-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 94/05550
FILING DATE: 05-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Weiser, Gerard J.
REGISTRATION NUMBER: 19,763
REFERENCE/DOCKET NUMBER: 989.6411P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-875-8383
TELEFAX: 215-875-8394
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 426 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-737-248-4

```

```

Query Match 65.9%; Score 1219.5; DB 3; Length 426;
Best Local Similarity 97.4%; Pred. No. 2.8e-104;
Matches 229; Conservative 2; Mismatches 3; Indels 1; Gaps 1;
QY 1 MSPILGWYKIKGLVOPTRLLLEYLEEKYEHLVERDEGDKWRNKKFELGLEFNPPLYID 60
DB 1 MSPILGWYKIKGLVOPTRLLLEYLEEKYEHLVERDEGDKWRNKKFELGLEFNPPLYID 60
QY 61 GVKLTQSMALIRYIADKNHMLGCGPKERAELISMLGAVLDIRGVSRIVASDFETLKY 120
DB 61 GVKLTQSMALIRYIADKNHMLGCGPKERAELISMLGAVLDIRGVSRIVASDFETLKY 120
QY 121 DFLSKLPEMLKMEEDRLCHKTYLNGDVTNHPDMLYDALDVLVLYMPMCLDAPPKLVCFK 180
DB 121 DFLSKLPEMLKMEEDRLCHKTYLNGDVTNHPDMLYDALDVLVLYMPMCLDAPPKLVCFK 180
QY 121 DFLSKLPEMLKMEEDRLCHKTYLNGDVTNHPDMLYDALDVLVLYMPMCLDAPPKLVCFK 180
DB 121 DFLSKLPEMLKMEEDRLCHKTYLNGDVTNHPDMLYDALDVLVLYMPMCLDAPPKLVCFK 180
QY 181 KRLEAIPOIDKYLKSSKYIAMPLOGQATFGGDDHPKSDLVPRGS-SLPTICSSG 235
DB 181 KRLEAIPOIDKYLKSSKYIAMPLOGQATFGGDDHPKSDLVPRGS-SLPTICSSG 235

```

```

RESULT 14
US-08-395-507-2
Sequence 2, Application US/08395507
Patent No. 5578456
GENERAL INFORMATION:
APPLICANT: Fujimura, Katsuya
APPLICANT: Ueno, Eiichi
APPLICANT: Fujii, No. 5578456uyuki

```

```

APPLICANT: Okada, Masahisa
TITLE OF INVENTION: Anti-treponema Pallidum Antibody
TITLE OF INVENTION: Immunoassay
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: Pollock, Vande Sande & Priddy
STREET: 190 W Street, NW, Suite 800
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/395,507
FILING DATE: 27-FEB-1995
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP Hei-6-54672
FILING DATE: 28-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Amerinick, Burton A.
REGISTRATION NUMBER: 24,852
REFERENCE/DOCKET NUMBER: 1581/007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 331-7111
TELEFAX: (202) 293-6229
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 362 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: Protein
LOCATION: 1..362
OTHER INFORMATION: /note="G17 Antigen"
US-08-395-507-2

```

```

Query Match 65.9%; Score 1219; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 2.5e-104;
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSPILGWYKIKGLVOPTRLLLEYLEEKYEHLVERDEGDKWRNKKFELGLEFNPPLYID 60
DB 1 MSPILGWYKIKGLVOPTRLLLEYLEEKYEHLVERDEGDKWRNKKFELGLEFNPPLYID 60
QY 61 GVKLTQSMALIRYIADKNHMLGCGPKERAELISMLGAVLDIRGVSRIVASDFETLKY 120
DB 61 GVKLTQSMALIRYIADKNHMLGCGPKERAELISMLGAVLDIRGVSRIVASDFETLKY 120
QY 121 DFLSKLPEMLKMEEDRLCHKTYLNGDVTNHPDMLYDALDVLVLYMPMCLDAPPKLVCFK 180
DB 121 DFLSKLPEMLKMEEDRLCHKTYLNGDVTNHPDMLYDALDVLVLYMPMCLDAPPKLVCFK 180
QY 121 DFLSKLPEMLKMEEDRLCHKTYLNGDVTNHPDMLYDALDVLVLYMPMCLDAPPKLVCFK 180
DB 121 DFLSKLPEMLKMEEDRLCHKTYLNGDVTNHPDMLYDALDVLVLYMPMCLDAPPKLVCFK 180
QY 181 KRLEAIPOIDKYLKSSKYIAMPLOGQATFGGDDHPKSDLVPRGS 226
DB 181 KRLEAIPOIDKYLKSSKYIAMPLOGQATFGGDDHPKSDLVPRGS 226

```

```

RESULT 15
US-08-102-757-9
Sequence 9, Application US/08102757
Patent No. 5480864
GENERAL INFORMATION:
APPLICANT: Tajima, Masahiro

```

APPLICANT: Ohnuma, Manami
APPLICANT: Lerner, Ethan A.
TITLE OF INVENTION: Modified Maxadilan Protein, Its
TITLE OF INVENTION: Preparation and Use, and DNA encoding the Protein
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flunegan, Henderson, Farabow, Garrett &
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/102,757
FILING DATE: 01-JAN-1993
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 05136.0003-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 291 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-102-757-9

Query Match 65.2%; Score 1207; DB 1; Length 291;
Best Local Similarity 99.1%; Pred. No. 2,4e-103;
Matches 224; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MSPIIGYWKIKGIYOPTRLILEYLEEKYEHLVERDEGDKWRNKKFELGLEFPNLPYYID 60
DB 1 MSPIIGYWKIKGIYOPTRLILEYLEEKYEHLVERDEGDKWLNKKFELGLEFPNLPYYID 60
QY 61 GYVKLTQSMATIRYTADKNMMLGGCPKERAETSMLEGAVIDIRYGSRIAYSKDFETLKV 120
DB 61 GYVKLTQSMATIRYTADKNMMLGGCPKERAETSMLEGAVIDIRYGSRIAYSKDFETLKV 120
QY 121 DFLSKLPMLKMFEDRLCHKTYLNGDHTVPDFMLYDALDVVLYMDPMLDAFPKLVCFK 180
DB 121 DFLSKLPMLKMFEDRLCHKTYLNGDHTVPDFMLYDALDVVLYMDPMLDAFPKLVCFK 180
QY 181 KRLEAIPOIDKTYLKSKSYTAMPLOGWQATFGGDDHPPKSDLVPRGS 226
DB 181 KRLEAIPOIDKTYLKSKSYTAMPLOGWQATFGGDDHPPKSDLVPRGS 226

Search completed: March 18, 2001, 05:30:32
Job time: 3401 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2001, 04:58:12 ; Search time 66.39 Seconds
(without alignments)
350.805 Million cell updates/sec

Title: US-09-402-488A-2
Perfect score: 1650
Sequence: 1 MSPILGYWKIKGLVQPTRL.....PKQSHNDGDEEIPPEXLYQ 343

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 6790655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1170	63.2	219 2	A26484 glutathione transf
2	990	53.5	218 2	A45556 glutathione S-tran
3	927	50.1	218 2	A45523 glutathione transf
4	654	35.4	209 2	A48388 glutathione S-tran
5	483.5	26.1	218 2	S33860 glutathione transf
6	483.5	26.0	218 2	A29794 glutathione transf
7	480.5	26.0	218 2	S32425 glutathione transf
8	480.5	26.0	218 2	A47486 glutathione transf
9	478.5	25.9	218 2	S65674 glutathione transf
10	474.5	25.6	218 2	A39375 glutathione transf
11	473.5	25.6	217 2	JX0095 glutathione transf
12	470.5	25.4	218 2	S13202 glutathione transf
13	469.5	25.4	218 2	A23732 glutathione transf
14	468.5	25.3	218 2	B34159 glutathione transf
15	467.5	25.3	218 2	S01719 glutathione transf
16	464.5	25.1	218 2	B29231 glutathione transf
17	464.5	25.1	218 2	A46048 glutathione transf
18	462.5	25.0	218 2	A29036 glutathione transf
19	456.5	24.7	218 1	XURG4 glutathione transf
20	453.5	24.5	218 2	A46143 mu-class glutathio
21	453.5	24.5	218 2	B28846 glutathione transf
22	449.5	24.3	225 2	A35295 glutathione transf
23	429.5	23.2	220 2	S18464 glutathione transf
24	385.5	20.8	219 2	S50146 major allergen Dpl
25	367	19.8	65 1	HULXH thrombin inhibitor
26	359	19.4	65 2	S78521 hirudin IIT - me
27	355	19.2	65 2	S05676 hirudin IIT - me
28	353	19.1	65 2	S78520 hirudin IITa - me
29	352	19.0	65 2	S05677 hirudin IITb - me

30	352	19.0	65 2	S05675 hirudin IITa - med
31	349	18.9	65 2	S05674 hirudin IITb - med
32	346	18.7	65 2	S05673 hirudin IITa - med
33	344	18.6	65 2	S05678 hirudin IIT - medic
34	339	18.3	65 2	S05679 hirudin IIT - medic
35	329	17.8	72 2	A37417 thrombin inhibitor
36	311	16.8	66 2	A24350 thrombin inhibitor
37	295	15.9	55 2	S05672 glutathione transf
38	257.5	13.9	142 2	S17462 glutathione transf
39	249	13.5	63 2	A53883 glutathione transf
40	248.5	13.4	208 2	S41933 glutathione transf
41	247	13.4	84 2	S33329 hirudin HM2 - leec
42	244	13.2	210 2	S71958 glutathione transf
43	242	13.1	203 2	S13780 glutathione transf
44	240	13.0	210 1	A37378 glutathione transf
45	238	12.9	210 2	S71957 glutathione transf

ALIGNMENTS

RESULT 1
A26484
glutathione transferase (EC 2.5.1.18) - fluke (Schistosoma japonicum) (fragment)
C:Species: Schistosoma japonicum
C:Date: 25-Oct-1987 #sequence_revision 30-Sep-1989 #text_change 03-Feb-1994
C:Accession: A94139; A26484; A28315
R:Smith, D.B.; Davern, K.M.; Board, P.G.; Tiu, W.U.; Garcia, E.G.; Mitchell, G.F.
Proc. Natl. Acad. Sci. U.S.A. 83, 8703-8707, 1986
A:Title: Mr 26,000 antigen of Schistosoma japonicum recognized by resistant WEHI 129/
A:Reference number: A94139; MID:87041520
A:Accession: A94139
A:Molecule type: mRNA
A:Residues: 1-219 <SMI>
R:Smith, D.B.; Davern, K.M.; Board, P.G.; Tiu, W.U.; Garcia, E.G.; Mitchell, G.F.
Proc. Natl. Acad. Sci. U.S.A. 84, 6541, 1987
A:Reference number: A94181
A:Contents: annotation; revision to residues 210-219
C:Superfamily: glutathione transferase
C:Keywords: transferase

Query Match 63.2% Score 1170; DB 2; Length 219;
Best Local Similarity 100.0%; Pred. No. 7.4e-89;
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PILGYWKIKGLVQPTRLLEYLEEKYEHLXERDEGKMKRKKFELERPNLPYIDGD 62
DB 4 PILGYWKIKGLVQPTRLLEYLEEKYEHLXERDEGKMKRKKFELERPNLPYIDGD 63

QY 63 VKLTQSMALIRYIADKNNMLGCCPKERAETSMLEGAVIDIRYGVSRATYASKDFETLKVDF 122
DB 64 VKLTQSMALIRYIADKNNMLGCCPKERAETSMLEGAVIDIRYGVSRATYASKDFETLKVDF 123

QY 123 LSKLPEMLKMFEDRLCKHTYLNDSVTHPPDFMLYDALDVLYMDPKLDAFPKLVCEKKR 182
DB 124 LSKLPEMLKMFEDRLCKHTYLNDSVTHPPDFMLYDALDVLYMDPKLDAFPKLVCEKKR 183

QY 183 IEALPQIDKYLKSSKRYLAWPLQGWATFGGDDHPK 218
DB 184 IEALPQIDKYLKSSKRYLAWPLQGWATFGGDDHPK 219

RESULT 2
A45556
glutathione S-transferase - fluke (Schistosoma mansoni)
C:Species: Schistosoma mansoni
C:Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 03-May-1996
C:Accession: A45556
R:Wright, M.D.; Harrison, R.A.; Melder, A.M.; Newport, G.R.; Mitchell, G.F.
Mol. Biochem. Parasitol. 49, 177-179, 1991
A:Title: Another 26-kilodalton glutathione S-transferase of Schistosoma mansoni.
A:Reference number: A45556; MID:92131046

A.Contents: clone pGT 875
 A.Accession: A28946
 A.Molecule type: mRNA
 A.Residues: 2-218 <PEA>
 A.Cross-references: GB:J03952; NID:9193687; PIDN:AAA37747.1; PID:9309278
 J.Pearson, W.R.; Windle, J.J.; Morrow, J.F.; Benson, A.M.; Tatalay, P.
 J. Biol. Chem. 258, 2052-2062, 1983
 A.Title: Increased synthesis of glutathione S-transferases in response to anticarcinogen
 A.Reference number: A92411; MUID:83109018
 A.Accession: A20831
 A.Molecule type: protein
 A.Residues: 2-7, 'X', 8-13, 'X', 15-23, 'X', 25, 'XX', 28, 'N', 30-31, 'X', 33, 'X', 35-39, 'W', 41 <PEI
 R.Hayes, J.D.; Kerr, L.A.; Peacock, S.D.; Cronshaw, A.D.; Mclellan, L.I.
 Biochem. J. 277, 501-512, 1991
 A.Title: Hepatic glutathione S-transferases in mice fed on a diet containing the anticar
 adient elution of the glutathione-Sepharose affinity matrix.
 A.Reference number: A24735; MUID:86042634
 A.Accession: 124735
 A.Molecule type: protein
 A.Residues: 2-7, 'X', 8-13, 'X', 15-23, 'X', 25 <MAN>
 R.Hayes, J.D.; Kerr, L.A.; Peacock, S.D.; Cronshaw, A.D.; Mclellan, L.I.
 Biochem. J. 277, 501-512, 1991
 A.Title: Hepatic glutathione S-transferases in mice fed on a diet containing the anticar
 adient elution of the glutathione-Sepharose affinity matrix.
 A.Reference number: S11311; MUID:96268466
 A.Accession: S16933; MUID:91315425
 A.Molecule type: protein
 A.Residues: 2-36 <NAV>
 R.Fernandes, C.L.; Dong, J.H.; Roebuck, B.D.; Chisari, F.V.; Montali, J.A.; Schmidt Jr.,
 Arch. Biochem. Biophys. 331, 104-116, 1996
 A.Title: Elevations of hepatic quinine reductase, glutathione, and alpha- and mu-class g
 A.Reference number: S11311; MUID:96268466
 A.Accession: S16933; MUID:91315425
 A.Molecule type: protein
 A.Residues: 97-108 <FER>
 A.Gene: GSTM1
 A.Introns: 12/3; 38/1; 59/3; 87/1; 120/3; 152/3; 189/3
 C:Superfamily: glutathione transferase
 C:Keywords: dimer; transferase
 F:2-218/Product: glutathione transferase mul #status experimental <MAT>

Query Match 26.1%; Score 483.5; DB 2: Length 218;
 Best Local Similarity 44.2%; Pred. No. 1.8e-32;
 Matches 92; Conservative 38; Mismatches 73; Indels 5; Gaps 1;

1 MSPILGWMKIKGLVPTLLLEYLEEKYEHLHYERDEG-----DKRNKKFELGLEFPNL 55
 1 MMLILGYMWRKLTPIRRLLEETDSSYDEKRTMGDADFDPSQWLNKFKLGLDFPNL 60
 56 PYIDGDKVLKTSMAIIRYIADKHNMLGCGPKRAEISMEGAVLDIRYGVSRVATSKDF 115
 61 PYLDGSRKITSMAIRYLARKHNLHLDGETEERIRADIVENQVMDTRMQLMLCYNPDF 120
 116 EFLKVDLSLPEMLKMFEDRLCHKTYLNGDHVTHPEMLYDALDVLYLMDPCGLDAFPP 175
 121 EKQKEPFLKTIPEKMKLYSEFLGKRWFGADKVTYVDLAYDILQYRMFEKCLDAFPN 180
 176 LWCFFKRIEAIPOIDKYLKSSKYIAMPL 203
 181 LDFLAREFGLKRIKISYMKSSRYIATPI 208

RESULT 6
 A29794
 glutathione transferase (EC 2.5.1.18) class mu chain 3 - rat
 N:Alternate names: glutathione S-transferase Yb1; ligandin
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 28-Aug-1989 #sequence-revision 28-Aug-1989 #text-change 18-Jun-1999
 C:Accession: A29794; A25510; A24085; B61004; F24735; A33397; S27111; S17167; A26187
 J.Chang, C.; Saltzman, A.G.; Sorensen, N.S.; Hlipakka, R.A.; Liao, S.
 J. Biol. Chem. 262, 11901-11903, 1987
 A.Title: Identification of glutathione S-transferase Yb-1 mRNA as the androgen-repressed

A.Reference number: A29794; MUID:87308179
 A.Accession: A29794
 A.Molecule type: mRNA
 A.Residues: 1-218 <CH>
 A.Cross-references: GB:J02810; NID:9204514; PIDN:AAA1293.1; PID:9204515
 R.Lai, H.C.J.; Grove, G.; Tu, C.P.D.
 Nucleic Acids Res. 14, 6101-6114, 1986
 A>Title: Cloning and sequence analysis of a cDNA for a rat liver glutathione S-transf
 A.Reference number: A25510; MUID:86312882
 A.Accession: A25510
 A.Molecule type: mRNA
 A.Residues: 1-218 <LAT>
 A.Cross-references: GB:X04229; NID:956337; PIDN:CAA27811.1; PID:956338
 R.Ding, G.J.F.; Lu, A.Y.H.; Pickett, C.B.
 J. Biol. Chem. 260, 13268-13271, 1985
 A.Reference number: A24085; MUID:86033768
 A.Accession: A24085
 A.Molecule type: mRNA
 A.Residues: 1-198, 'NC', 201-218 <DIN>
 A.Cross-references: GB:M11719; NID:9204502; PIDN:AAA1287.1; PID:9204503
 A.Experimental source: clone pGTA/C44
 R.Chang, L.H.; Hsieh, J.C.; Chen, W.L.; Tam, M.F.
 Electrophoresis 11, 589-593, 1990
 A>Title: Identification of rat liver glutathione S-transferase Yb subunits by partial
 lal isoelectric focusing gel.
 A.Reference number: A61004; MUID:91031411
 A.Accession: B61004
 A.Molecule type: protein
 A.Residues: 2-26 <CH2>
 R.Mannervik, B.; Alin, P.; Guttenberg, C.; Jansson, H.; Tahiri, M.K.; Warholm, M.; Jor
 Proc. Natl. Acad. Sci. U.S.A. 82, 7202-7206, 1985
 A>Title: Identification of three classes of cytosolic glutathione transferase common
 A.Reference number: A24735; MUID:86042634
 A.Accession: F24735
 A.Molecule type: protein
 A.Residues: 2-20 <MAN>
 R.Hsieh, J.C.; Liu, L.F.; Chen, W.L.; Tam, M.F.
 Biochem. Biophys. Res. Commun. 162, 1147-1154, 1989
 A>Title: Expression of Yb-1 glutathione S-transferase using a baculovirus expression
 A.Reference number: A33397; MUID:89350924
 A.Accession: A33397
 A.Molecule type: protein
 A.Residues: 2-21, 212-218 <HS1>
 R.Katusz, R.M.; Bono, B.; Colman, R.F.
 Arch. Biochem. Biophys. 298, 667-677, 1992
 A>Title: Identification of Tyr(115) labeled by S-(4-Dromo-2,3-dioxobutyl)glutathione
 A.Reference number: S27111; MUID:93037509
 A.Accession: S27111
 A.Molecule type: protein
 A.Residues: 2-20; 83-86, 'X', 88-96; 109-115, 'X', 117-122 <KAT>
 R.Hsieh, J.C.; Huang, S.C.; Chen, W.L.; Lai, Y.C.; Tam, M.F.
 Biochem. J. 278, 293-297, 1991
 A>Title: Cysteine-86 is not needed for the enzymic activity of glutathione S-transfer
 A.Reference number: S17167; MUID:91354218
 A.Accession: S17167
 A:Status: preliminary
 A:Molecule type: protein
 A.Residues: 2-218 <HS2>
 C:Superfamily: glutathione transferase
 C:Keywords: dimer; transferase

Query Match 26.1%; Score 483.5; DB 2: Length 218;
 Best Local Similarity 44.2%; Pred. No. 1.8e-32;
 Matches 92; Conservative 38; Mismatches 73; Indels 5; Gaps 1;

1 MSPILGWMKIKGLVPTLLLEYLEEKYEHLHYERDEG-----DKRNKKFELGLEFPNL 55
 1 MMLILGYMWRKLTPIRRLLEETDSSYDEKRTMGDADFDPSQWLNKFKLGLDFPNL 60
 56 PYIDGDKVLKTSMAIIRYIADKHNMLGCGPKRAEISMEGAVLDIRYGVSRVATSKDF 115
 61 PYLDGSRKITSMAIRYLARKHNLHLDGETEERIRADIVENQVMDTRMQLMLCYNPDF 120

QY 116 ETLKVDLSKLEPEMKMEDRLCHKTYLNGDHPDMLVDALDVLVYMDPMLCLADAPK 175
| | | | | : | | : | : | : | : | : | | | | | : | | | | | : | | | | |
DB 121 EKQKEPFLKTPPEKRLKLSSEFLGKRPWFGDKVTVVDLADVLIDQVHFEKCLADAPN 180
| | | | | : | | : | : | : | : | : | | | | | : | | | | | : | | | | |
QY 176 LVCFKKRIEALPQIDKYLKSSKYIAMP 203
| | | | | : | | : | : | : | : | : | | | | | : | | | | | : | | | | |
DB 181 LKDFLAREGLKRLISAVYMKSSRYLSTPI 208
| | | | | : | | : | : | : | : | : | | | | | : | | | | | : | | | | |
RESULT 7
S32425
glutathione transferase (EC 2.5.1.18) class mu, GSTM4 (version 2) - human
N:Alternate names: glutathione transferase mu4
C:Species: Homo sapiens (man)
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 18-Jun-1999
R:Accession: S32425; S29337
R:Zhong, S.; Spurr, N.K.; Hayes, J.D.; Wolf, C.R.
Biochem. J. 291, 41-50, 1993
A:Title: Deduced amino acid sequence, gene structure and chromosomal location of a novel
A:Reference number: S32424; MUID:93228631
A:Accession: S32425
A:Molecule type: DNA
A:Residues: 1-218 <ZHO>
A:Cross-references: EMBL:X68677; NID:g31934; PIDN:CAA0637.1; PID:g31935
A:Note: the authors translated the codon AAG for residue 182 as Arg and CCA for residue
C:Genetics:
A:Gene: GDB:GSTM4
A:Cross-references: GDB:134191; OMIM:138333
A:Map position: 1p13.3-1p13.3
A:Introns: 12/3; 38/1; 59/3; 87/1; 120/3; 152/3; 189/3
C:Superfamily: glutathione transferase
C:Keywords: dimer; transferase

Query Match 26.0%; Score 480.5; DB 2; Length 218;
Best Local Similarity 44.0%; Pred. No. 3.3e-32;
Matches 95; Conservative 37; Mismatches 79; Indels 5; Gaps 1;
QY 1 MSPILGYKIKIGVQPTLLLEYLEEKEEHLHYEDG-----GCDKRRKKFELGIEPPL 55
| | | | | : | | : | : | : | : | : | | | | | : | | | | | : | | | | |
DB 1 MPMILGYMDIRGLAVALIRLLLETDSSYEKKTYMGADPDVDRSQMLNEKRLGIDPPL 60
| | | | | : | | : | : | : | : | : | | | | | : | | | | | : | | | | |
QY 56 PYIDGDKVLKQSMALIRYADKHNMLGCGPKERAELSMLEGAVIDIRGVSRIAYSKDF 115
| | | | | : | | : | : | : | : | : | | | | | : | | | | | : | | | | |
DB 61 PYLDGAHKRITQSNAILCYIARKHMLCGETEPEEKIRVDILENQADVSNQALRVCYSPDF 120
| | | | | : | | : | : | : | : | : | | | | | : | | | | | : | | | | |
QY 116 ETLKVDLSKLEPEMKMEDRLCHKTYLNGDHPDMLVDALDVLVYMDPMLCLADAPK 175
| | | | | : | | : | : | : | : | : | | | | | : | | | | | : | | | | |
DB 121 EKLKEPEYLEELPTMQHRSQFLGKRPWFVGDKITFVDLAVDLDLHRIFEPCIDAPFN 180
| | | | | : | | : | : | : | : | : | | | | | : | | | | | : | | | | |
QY 176 LVCFKKRIEALPQIDKYLKSSKYIAMPLOGMQATFG 211
| | | | | : | | : | : | : | : | : | | | | | : | | | | | : | | | | |
DB 181 LKDFISREGELEKISAVYMKSSRYLSTPIYTRVAWVG 216
| | | | | : | | : | : | : | : | : | | | | | : | | | | | : | | | | |
RESULT 8
A47486
glutathione transferase (EC 2.5.1.18) class mu, GSTM4 (version 1) - human
C:Species: Homo sapiens (man)
C:Date: 02-Jun-1995 #sequence_revision 02-Jun-1995 #text_change 18-Jun-1999
R:Accession: A47486; S36782; I37438; S45685
R:Comstock, K.E.; Johnson, K.J.; Rifkenbery, D.; Henner, W.D.
J. Biol. Chem. 268, 16958-16965, 1993
A:Title: Isolation and analysis of the gene and cDNA for a human mu class glutathione S-
A:Reference number: A47486; MUID:93352467
A:Accession: A47486
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-218 <COM>
A:Cross-references: GB:M96233
A:Accession: B47486
A:Status: preliminary

A:Molecule type: mRNA
A:Residues: 1-218 <CO2>
A:Cross-references: GB:M96234; NID:g306818; PIDN:AAA57347.1; PID:g306819
A:Experimental source: HeLa cells
R:Ross, V.L.; Board, P.G.
Biochem. J. 294, 373-380, 1993
A:Title: Molecular cloning and heterologous expression of an alternatively spliced hu
A:Reference number: S36782; MUID:93384505
A:Accession: S36782
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-218 <ROS>
A:Cross-references: EMBL:M99422
A:Experimental source: testis
R:Comstock, K.E.; Widersten, M.; Hao, X.Y.; Henner, W.D.; Mannervik, B.
Arch. Biochem. Biophys. 311, 487-495, 1994
A:Title: A comparison of the enzymatic and physicochemical properties of human glutat
A:Reference number: S45684; MUID:94263230
A:Contents: annotation
R:Taylor, J.B.; Oliver, J.; Sherrington, R.; Pemble, S.E.
Biochem. J. 274, 587-593, 1991
A:Title: Structure of human glutathione S-transferase class Mu genes.
A:Reference number: I37438; MUID:91174774
A:Accession: I37438
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 39-120 <RES>
A:Cross-references: EMBL:X56837; NID:g31936; PIDN:CAA0167.1; PID:g31937
C:Genetics:
A:Gene: GDB:GSTM4
A:Cross-references: GDB:134191; OMIM:138333
A:Map position: 1p13.3-1p13.3
A:Introns: 12/3; 38/1; 59/3; 87/1; 120/3; 152/3; 189/3
C:Superfamily: glutathione transferase
C:Keywords: alternative splicing; dimer; transferase

Query Match 26.0%; Score 480.5; DB 2; Length 218;
Best Local Similarity 44.0%; Pred. No. 3.3e-32;
Matches 95; Conservative 37; Mismatches 79; Indels 5; Gaps 1;
QY 1 MSPILGYKIKIGVQPTLLLEYLEEKEEHLHYEDG-----DKRRKKFELGIEPPL 55
| | | | | : | | : | : | : | : | : | | | | | : | | | | | : | | | | |
DB 1 MPMILGYMDIRGLAVALIRLLLETDSSYEKKTYMGADPDVDRSQMLNEKRLGIDPPL 60
| | | | | : | | : | : | : | : | : | | | | | : | | | | | : | | | | |
QY 56 PYIDGDKVLKQSMALIRYADKHNMLGCGPKERAELSMLEGAVIDIRGVSRIAYSKDF 115
| | | | | : | | : | : | : | : | : | | | | | : | | | | | : | | | | |
DB 61 PYLDGAHKRITQSNAILCYIARKHMLCGETEPEEKIRVDILENQADVSNQALRVCYSPDF 120
| | | | | : | | : | : | : | : | : | | | | | : | | | | | : | | | | |
QY 116 ETLKVDLSKLEPEMKMEDRLCHKTYLNGDHPDMLVDALDVLVYMDPMLCLADAPK 175
| | | | | : | | : | : | : | : | : | | | | | : | | | | | : | | | | |
DB 121 EKLKEPEYLEELPTMQHRSQFLGKRPWFVGDKITFVDLAVDLDLHRIFEPCIDAPFN 180
| | | | | : | | : | : | : | : | : | | | | | : | | | | | : | | | | |
QY 176 LVCFKKRIEALPQIDKYLKSSKYIAMPLOGMQATFG 211
| | | | | : | | : | : | : | : | : | | | | | : | | | | | : | | | | |
DB 181 LKDFISREGELEKISAVYMKSSRYLSTPIYTRVAWVG 216
| | | | | : | | : | : | : | : | : | | | | | : | | | | | : | | | | |
RESULT 9
S65674
glutathione transferase (EC 2.5.1.18) class mu chain 7.8 - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 18-Jun-1999
R:Accession: S65674; S30380
R:Lee, S.H.; Lee, Y.S.; Han, J.S.; Kim, Y.S.; Koh, J.K.
Arch. Biochem. Biophys. 318, 424-429, 1995
A:Title: Cloning and expression of a cDNA for Mu-Class Glutathione S-transferase from
A:Reference number: S65674; MUID:95251394
A:Accession: S65674
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-218 <LEE>
A:Cross-references: EMBL:L23766; NID:g388198; PIDN:AAA69665.1; PID:g388199

R:Primiano, T.; Novak, R.F.
Arch. Biochem. Biophys. 301, 404-410, 1993
A:Title: Purification and characterization of class mu glutathione S-transferase isozyme
A:Reference number: S30380; MUID:93213177
A:Accession: S30380
A:Molecule type: protein
A:Residues: 2-21 <PRI>
C:Superfamily: glutathione transferase
C:Keywords: dimer; transferase

Query Match 25.9%; Score 478.5; DB 2; Length 218;
Best Local Similarity 44.2%; Pred. No. 4.8e-32;
Matches 96; Conservative 36; Mismatches 80; Indels 5; Gaps 1;
QY 1 MSPILGYWKIKGLVOPRLLEYLEEKEEHLERDEG-----DKMNRKFEGLPEPNTL 55
1 |MPTLGYWNRGLALPRLMLEETDTSYEKKYTMGDAPYDRSQWLNKFKLGLDEPNL 60
DB 1 |MPTLGYWNRGLALPRLMLEETDTSYEKKYTMGDAPYDRSQWLNKFKLGLDEPNL 60
QY 56 PYIYDGVKLTQSMATIRYIADKHNMLGGCPKRAEISMLEGAVLDIRYGSRIATYSKDF 115
1 |MPTLGYWNRGLALPRLMLEETDTSYEKKYTMGDAPYDRSQWLNKFKLGLDEPNL 60
DB 61 PYLIDGTHKLTQSNALIRYIARKHNLGCGEVEERIRVDILENQLMDNRFLQVNVVCSPDF 120
QY 116 ETLKVDLSKLPEMLKFEEDRLCHKTYLNGDHTHPDMLYDALDVLVYMDPCLDAFPK 175
1 |MPTLGYWNRGLALPRLMLEETDTSYEKKYTMGDAPYDRSQWLNKFKLGLDEPNL 60
DB 121 ETLKVDLSKLPEMLKFEEDRLCHKTYLNGDHTHPDMLYDALDVLVYMDPCLDAFPN 180
QY 176 LVCFKRRIEAIPOIDKYLKSSKIYAMPLOQWATFG 212
1 |MPTLGYWNRGLALPRLMLEETDTSYEKKYTMGDAPYDRSQWLNKFKLGLDEPNL 60
DB 181 LKDFHVRFEGLPKISAYMKSSRFLRPVFLKATWTG 217

RESULT 10
A39375

glutathione transferase (EC 2.5.1.18) mu-2 [validated] - human
N:Alternate names: glutathione S-transferase GST4, muscle; glutathione transferase GST6;
C:Species: Homo sapiens (man)
C:Date: 20-Mar-1992 #sequence_revision 06-Jan-1995 #text_change 21-Jul-2000
C:Accession: A39375; A28810; C35187; S13305; S14344
R:Vorachek, W.R.; Pearson, W.R.; Rule, G.S.
Proc. Natl. Acad. Sci. U.S.A. 88, 4443-4447, 1991
A:Title: Cloning, expression, and characterization of a class-mu glutathione transferase
A:Reference number: A39375; MUID:91239584
A:Accession: A39375
A:Molecule type: mRNA
A:Residues: 1-218 <WOR>
A:Cross-references: GB:M63509; NID:q183300; PIDN:AAA0963.1; PID:q183301
A:Experimental source: myoblast
A:Note: Translation of Initiator Met is not shown
R:Singh, S.V.; Ahmad, H.; Kurosky, A.; Awasthi, Y.C.
Arch. Biochem. Biophys. 264, 13-22, 1988
A:Title: Purification and characterization of unique glutathione S-transferases from hum
A:Reference number: A90081; MUID:88280250
A:Accession: A28810
A:Molecule type: protein
A:Residues: 2-7, 'X', 9-10, 'X', 12-14 <SIN>
A:Experimental source: skeletal muscle
R:Tsuchida, S.; Maki, T.; Sato, K.
J. Biol. Chem. 265, 7150-7157, 1990
A:Title: Purification and characterization of glutathione transferases with an activity
A:Reference number: A35187; MUID:90237002
A:Accession: C35187
A:Molecule type: protein
A:Residues: 2-10, 'X', 12-14, 'X', 16-17, 'X', 19-21, 'XXXX' <TSU>
A:Experimental source: heart
R:Hussey, A.J.; Kerr, B.A.; Cromshaw, A.D.; Harrison, D.J.; Hayes, J.D.
Biochem. J. 273, 323-332, 1991
A:Title: Variation in the expression of mu-class glutathione S-transferase isoenzymes fr
A:Reference number: S13305; MUID:91119559
A:Accession: S13305
A:Status: preliminary
A:Molecule type: protein
A:Residues: 2-25 <HDS>

A:Experimental source: skeletal muscle
R:Suzuki, T.; Shaw, D.C.; Board, P.G.
Biochem. J. 274, 405-408, 1991
A:Title: Purification and characterization of acidic glutathione S-transferase 6 from
A:Reference number: S14344; MUID:9114747
A:Accession: S14344

A:Molecule type: protein
A:Residues: 2-29, 'K', 31 <SUZ>
A:Experimental source: brain
C:Genetics:
A:Gene: GDB:GSTW2; GST4
A:Cross-references: GDB:128593; OMIM:138380
A:Map position: 1p13.3-1p13.3
C:Superfamily: glutathione transferase
C:Keywords: brain; cardiac muscle; heart; homodimer; skeletal muscle; transferase
F:2-218/Product: glutathione transferase mu-2 #status experimental <MAT>

Query Match 25.6%; Score 474.5; DB 2; Length 218;
Best Local Similarity 44.0%; Pred. No. 1e-31;
Matches 95; Conservative 35; Mismatches 81; Indels 5; Gaps 1;

QY 1 MSPILGYWKIKGLVOPRLLEYLEEKEEHLERDEG-----DKMNRKFEGLPEPNTL 55
1 |MPTLGYWNRGLALPRLMLEETDTSYEKKYTMGDAPYDRSQWLNKFKLGLDEPNL 60
DB 1 |MPTLGYWNRGLALPRLMLEETDTSYEKKYTMGDAPYDRSQWLNKFKLGLDEPNL 60
QY 56 PYIYDGVKLTQSMATIRYIADKHNMLGGCPKRAEISMLEGAVLDIRYGSRIATYSKDF 115
1 |MPTLGYWNRGLALPRLMLEETDTSYEKKYTMGDAPYDRSQWLNKFKLGLDEPNL 60
DB 61 PYLIDGTHKLTQSNALIRYIARKHNLGCGEVEERIRVDILENQLMDNRFLQVNVVCSPDF 120
QY 116 ETLKVDLSKLPEMLKFEEDRLCHKTYLNGDHTHPDMLYDALDVLVYMDPCLDAFPK 175
1 |MPTLGYWNRGLALPRLMLEETDTSYEKKYTMGDAPYDRSQWLNKFKLGLDEPNL 60
DB 121 ETLKVDLSKLPEMLKFEEDRLCHKTYLNGDHTHPDMLYDALDVLVYMDPCLDAFPN 180
QY 176 LVCFKRRIEAIPOIDKYLKSSKIYAMPLOQWATFG 211
1 |MPTLGYWNRGLALPRLMLEETDTSYEKKYTMGDAPYDRSQWLNKFKLGLDEPNL 60
DB 181 LKDFHVRFEGLPKISAYMKSSRFLRPVFLKATWTG 216

RESULT 11
JX0095

glutathione transferase (EC 2.5.1.18) b - guinea pig
N:Alternate names: glutathione S-alkyltransferase; glutathione S-aryltransferase; glu
C:Species: Cavia porcellus (guinea pig)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 18-Jun-1993
C:Accession: JX0095
R:Kamel, K.; Oshino, R.; Hara, S.
J. Biochem. 107, 111-117, 1990
A:Title: Amino acid sequence of glutathione S-transferase b from guinea pig liver.
A:Reference number: JX0095; MUID:90236961.
A:Accession: JX0095
A:Molecule type: protein
A:Residues: 1-217 <RAM>
C:Comment: Glutathione transferases are a family of enzymes that are multifunctional
C:Superfamily: glutathione transferase
C:Keywords: transferase

Query Match 25.6%; Score 473.5; DB 2; Length 217;
Best Local Similarity 45.1%; Pred. No. 1.2e-31;
Matches 92; Conservative 34; Mismatches 73; Indels 5; Gaps 1;

QY 5 LGYWKIKGLVOPRLLEYLEEKEEHLERDEG-----DKMNRKFEGLPEPNTL 59
1 |MPTLGYWNRGLALPRLMLEETDTSYEKKYTMGDAPYDRSQWLNKFKLGLDEPNL 60
DB 4 LGYWKIKGLVOPRLLEYLEEKEEHLERDEG-----DKMNRKFEGLPEPNTL 59
QY 60 DGVKLTQSMATIRYIADKHNMLGGCPKRAEISMLEGAVLDIRYGSRIATYSKDF 119
1 |MPTLGYWNRGLALPRLMLEETDTSYEKKYTMGDAPYDRSQWLNKFKLGLDEPNL 60
DB 64 DGVKLTQSMATIRYIADKHNMLGGCPKRAEISMLEGAVLDIRYGSRIATYSKDF 123
1 |MPTLGYWNRGLALPRLMLEETDTSYEKKYTMGDAPYDRSQWLNKFKLGLDEPNL 60
QY 120 VDFLSKLPEMLKFEEDRLCHKTYLNGDHTHPDMLYDALDVLVYMDPCLDAFPK 179
1 |MPTLGYWNRGLALPRLMLEETDTSYEKKYTMGDAPYDRSQWLNKFKLGLDEPNL 60
DB 124 AEFLEGPDKMKLFESQFLKLPMPFAGNKLTVYDVLAVDQYRMLPCKLEAFPNL 183

OY 180 KKRIEAIPOIDIKYLSSKYIAMPL 203
DB 184 ISREFGLEKISSYKSSRFLPKPL 207

RESULT 12

glutathione transferase (EC 2.5.1.18) Y1 - Chinese hamster
C:Species: Cricetus griseus (Chinese hamster)
C:Date: 18-Feb-1994 #sequence_revision 06-Jan-1995 #text_change 18-Jun-1999
C:Accession: S13202
R:Robert de Saint Vincent, B.; Hyrien, O.; Debatisse, M.; Buttin, G.
A:Title: Coamplification of mu class glutathione S-transferase genes and an adenylate de
A:Reference number: S13202; MUID:91031445
A:Accession: S13202
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-218 <ROB>
A:Cross-references: GB:X57489; NID:949541; PIDN:CAA0726.1; PID:949542
A:Note: in Genbank entry CUI1, release 109, the source is designated as Cricetus longi
C:Superfamily: glutathione transferase
C:Keywords: dimer; transferase

Query Match 25.4%; Score 470.5; DB 2; Length 218;

Best Local Similarity 43.3%; Pred. No. 2,2e-31;

Matches 90; Conservative 36; Mismatches 77; Indels 5; Gaps 1;

OY 1 MSPILGYWKIKGLVPTRLLEYLEEKYEHLIERDEG-----DKWRNKKFELGLEFPNL 55
DB 1 MPMILGYWVWVRLFTNPIRLLEYLETDSYEEKYTMGDAPDSRQWLNKFKLGLDFPNL 60
OY 56 PYIDGDKVLQSMALIRIADKHNMLGGCPKERAISMLEGAVDIRGVSRIAYSDF 115
DB 61 PYLDGSHKITYOSNALIRLARKHNLGCTEERIRVDIVENQAMDTRQMLMLCPDPF 120
OY 116 ETLKVDFLSKLPEMLKMFEDRLCHTKTYLNGDVHTPDPMLYDALDVLYMPDMLDAPFK 175
DB 121 EKQKEPEFLKTIPEKKMYSEFLGKRPWFAGDKVTLGFLAYVDLYQYQFEKCLDPFN 180
OY 176 LWCFKKRIEAIPOIDIKYLSSKYIAMPL 203
DB 181 LKDFLAREGLKISSAYMKTSLRFLRPI 208

RESULT 13

A23732
glutathione transferase (EC 2.5.1.18) mu - golden hamster

C:Species: Mesocricetus auratus (golden hamster)

C:Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 18-Jun-1999

C:Accession: A23732; S21908

R:Norris, J.S.; Schwartz, D.A.; Macleod, S.L.; Fan, W.; O'Brien, T.J.; Harris, S.E.; Tri
Mol. Endocrinol. 5: 979-986, 1991

A:Title: Cloning of a mu-class glutathione S-transferase complementary DNA and character

A:Reference number: A23732; MUID:92049380

A:Accession: A23732

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-218 <NOR>

A:Cross-references: EMBL:X61033; NID:949638; PIDN:CAA43368.1; PID:949639

A:Note: submitted to the EMBL Data Library, July 1991

C:Superfamily: glutathione transferase

C:Keywords: transferase

Query Match 25.4%; Score 469.5; DB 2; Length 218;

Best Local Similarity 44.3%; Pred. No. 2,6e-31;

Matches 90; Conservative 37; Mismatches 71; Indels 5; Gaps 1;

OY 1 MSPILGYWKIKGLVPTRLLEYLEEKYEHLIERDEG-----DKWRNKKFELGLEFPNL 55
DB 181 LKDFLAREGLKISSAYMKTSLRFLRPI 208

DB 1 MPTVLGYWDINRLAHAILRLLETTDTSYEEKYTMGDAPNDRSQWLNKFKLGLDFPNL 60

OY 56 PYIDGDKVLQSMALIRIADKHNMLGGCPKERAISMLEGAVDIRGVSRIAYSDF 115

DB 61 PYLDGSHKITYOSNALIRLARKHNLGCTEERIRVDIVENQAMDTRQMLMLCPDPF 120

OY 116 ETLKVDFLSKLPEMLKMFEDRLCHTKTYLNGDVHTPDPMLYDALDVLYMPDMLDAPFK 175

DB 121 EKQKEPEFLKTIPEKKMYSEFLGKRPWFAGDKVTLGFLAYVDLYQYQFEKCLDPFN 180

OY 176 LWCFKKRIEAIPOIDIKYLSSKY 198

DB 181 LKDFLAREGLKISSAYMKTSLRFLRPI 203

RESULT 14

B34159
glutathione transferase (EC 2.5.1.18) mu2 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 11-Apr-1990 #sequence_revision 25-Sep-1992 #text_change 18-Jun-1999

C:Accession: B34159; S30369; S30370; S30372

R:Tomson, A.T.; Goldsmith, M.E.; Pickett, C.B.; Cowan, K.H.

J. Biol. Chem. 264: 21582-21590, 1989

A:Title: Isolation, characterization, and expression in Escherichia coli of two murin

A:Reference number: A34159; MUID:90094327

A:Accession: B34159

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-218 <TON>

A:Cross-references: GB:J04666; NID:9193549; PIDN:AAA37706.1; PID:9309261

A:Note: this form was identified by PI 6.7

A:Accession: S30370

A:Molecule type: protein

A:Note: this form was identified by PI 7.9

A:Residues: 2-25 <AM2>

A:Accession: S30372

A:Molecule type: protein

A:Residues: 2-25 <AM3>

A:Note: this form was identified by PI 8.5

C:Superfamily: glutathione transferase

C:Keywords: transferase

Query Match 25.3%; Score 468.5; DB 2; Length 218;

Best Local Similarity 42.8%; Pred. No. 3,2e-31;

Matches 89; Conservative 41; Mismatches 73; Indels 5; Gaps 1;

OY 1 MSPILGYWKIKGLVPTRLLEYLEEKYEHLIERDEG-----DKWRNKKFELGLEFPNL 55

DB 1 MPTVLGYWDINRLAHAILRLLETTDTSYEEKYTMGDAPNDRSQWLNKFKLGLDFPNL 60

OY 56 PYIDGDKVLQSMALIRIADKHNMLGGCPKERAISMLEGAVDIRGVSRIAYSDF 115

DB 61 PYLDGSHKITYOSNALIRLARKHNLGCTEERIRVDIVENQAMDTRQMLMLCPDPF 120

OY 116 ETLKVDFLSKLPEMLKMFEDRLCHTKTYLNGDVHTPDPMLYDALDVLYMPDMLDAPFK 175

DB 121 EKQKEPEFLKTIPEKKMYSEFLGKRPWFAGDKVTLGFLAYVDLYQYQFEKCLDPFN 180

OY 176 LWCFKKRIEAIPOIDIKYLSSKYIAMPL 203

DB 181 LKDFLAREGLKISSAYMKTSLRFLRPI 208

RESULT 15

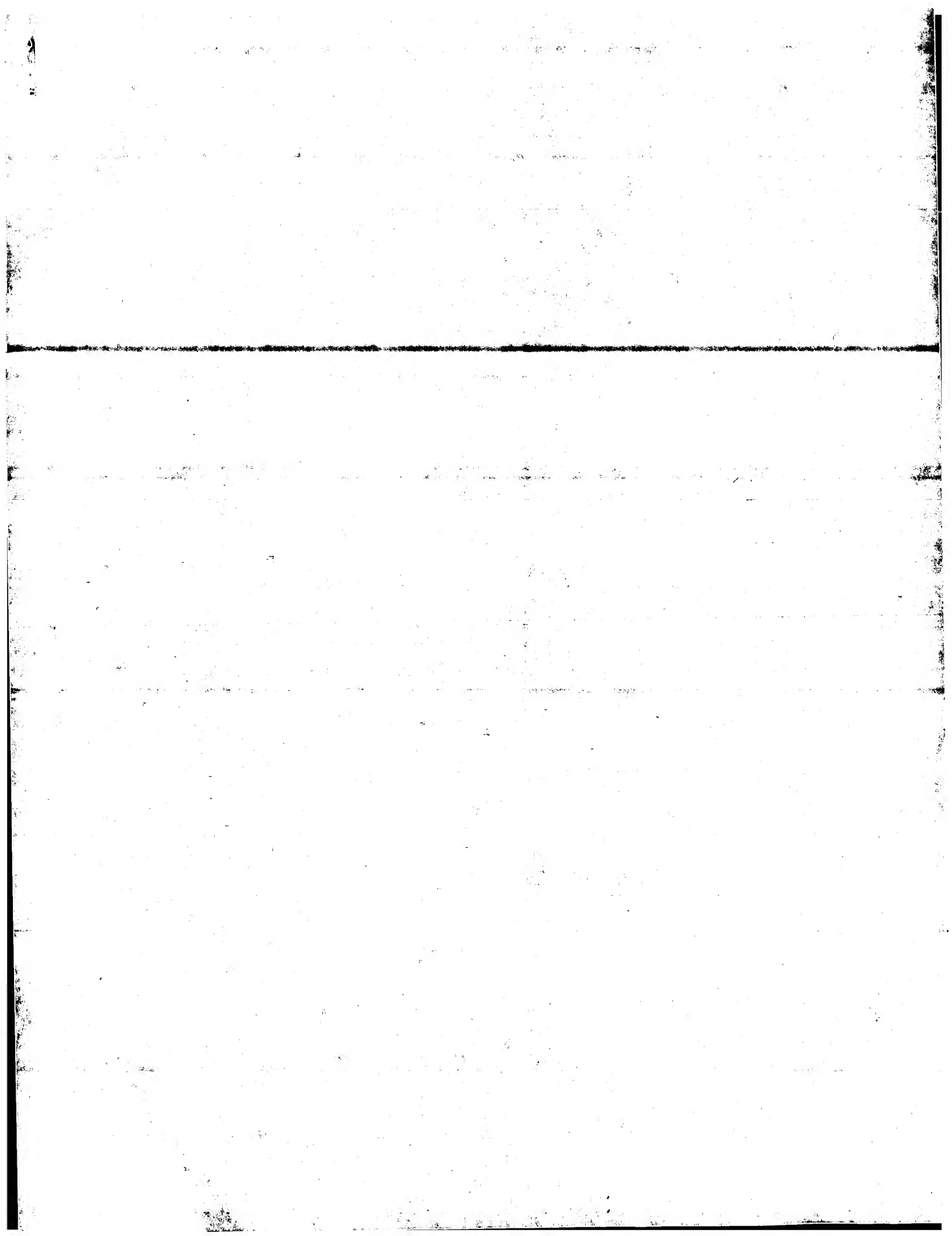
S01719

glutathione transferase (EC 2.5.1.18) class mu, GSTM1 - human
 N:Alternate names: glutathione S-transferase 1, mu (H-b); glutathione transferase 4; glu
 C:Species: Homo sapiens (man)
 C:Date: 30-Jun-1989 #sequence, revision 30-Jun-1989 #ext.change 21-Jul-2000
 C:Accession: S01719, A30770; S32424; S13905; B22457; D35187; S27188; I51867; E24
 R:DeJong, J.L., Chang, C.M., Whang-Peng, J., Knutsen, T., Tu, C.P.D.
 Nucleic Acids Res. 16, 8541-8554, 1988
 A:Title: The human liver glutathione S-transferase gene superfamily: expression and chro
 A:Reference number: S01719; MUID:88335606
 A:Contents: allele 1a (mu)
 A:Accession: S01719
 A:Molecule type: mRNA
 A:Residues: 1-218 <DEJ>
 A:Cross-references: EMBL:X08020; NID:g31923; PIDN:CAA30821.1; PID:g31924
 A:Experimental source: liver
 R:Seidgard, J.; Vorachek, W.R.; Pero, R.W.; Pearson, W.R.
 Proc. Natl. Acad. Sci. U.S.A. 85, 7293-7297, 1988
 A:Title: Hereditary differences in the expression of the human glutathione transferase a
 A:Reference number: A30770; MUID:89017184
 A:Contents: allele 1b (psi)
 A:Accession: A30770
 A:Molecule type: mRNA
 A:Residues: 1-172, 'N', 174-218 <SFI>
 A:Cross-references: EMBL:J03817; NID:g183668; PIDN:AAA59203.1; PID:g306812
 A:Experimental source: liver
 R:Chong, S.; Spurr, N.K.; Hayes, J.D.; Wolf, C.R.
 Biochem. J. 291, 41-50, 1993
 A:Title: Deduced amino acid sequence, gene structure and chromosomal location of a novel
 A:Reference number: S32424; MUID:93228651
 A:Accession: S32424
 A:Molecule type: DNA
 A:Residues: 1-43, 'T', 45-172, 'N', 174-218 <ZHO>
 A:Cross-references: EMBL:X68676
 A:Note: the authors translated the codon ACG for residue 44 as Ser
 R:Comstock, K.E.; Sanderson, B.J.; Clafin, G.; Henner, W.D.
 Nucleic Acids Res. 18, 3670, 1990
 A:Title: GST1 gene deletion determined by polymerase chain reaction.
 A:Reference number: I37437; MUID:90301515
 A:Accession: I37437
 A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMH
 A:Molecule type: DNA
 A:Residues: 60-118 <RES>
 A:Cross-references: EMBL:X51451; NID:g31922; PIDN:CAA35817.1; PID:g4378985
 R:Singhal, S.S.; Ahmad, H.; Sharma, R.; Gupta, S.; Haque, A.K.; Awasthi, Y.C.
 Arch. Biochem. Biophys. 285, 64-73, 1991
 A:Title: Purification and characterization of human muscle glutathione S-transferases: e
 A:Reference number: S13905; MUID:91119426
 A:Accession: S13905
 A:Molecule type: protein
 A:Residues: 2-15 <SIN>
 A:Experimental source: muscle
 R:Alin, P.; Mannervik, B.; Jornvall, H.
 FEBS Lett. 182, 319-322, 1985
 A:Title: Structural evidence for three different types of glutathione transferase in hum
 A:Reference number: A91336; MUID:85154554
 A:Accession: B22457
 A:Molecule type: protein
 A:Residues: 2-24 <ALI>
 R:Tsuchida, S.; Maki, T.; Sato, K.
 J. Biol. Chem. 265, 7150-7157, 1990
 A:Title: Purification and characterization of glutathione transferases with an activit
 A:Reference number: A35187; MUID:90237002
 A:Accession: D35187
 A:Molecule type: protein
 A:Residues: 2-10, 'X', 12-14, 'X', 16-17, 'X', 19-25 <TSU>
 A:Experimental source: heart
 R:Singhal, S.S.; Saxena, M.; Awasthi, S.; Ahmad, H.; Sharma, R.; Awasthi, Y.C.
 Biochim. Biophys. Acta 1171, 19-26, 1992
 A:Title: Gender related differences in the expression and characteristics of glutathione
 A:Reference number: S27188; MUID:93042004
 A:Accession: S27188
 A:Molecule type: protein
 A:Residues: 2-13 <STI>

A:Experimental source: colon
 R:Pearson, W.R.; Vorachek, W.R.; Xu, S.J.; Berger, R.; Hart, I.; Vannals, D.; Paters
 Am. J. Hum. Genet. 53, 220-233, 1993
 A:Title: Identification of class-mu glutathione transferase genes GSTM1-GSTM5 on huma
 A:Reference number: I51867; MUID:93304417
 A:Accession: I51867
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 125-172, 'N', 174-186 <RE2>
 A:Cross-references: GB:S62935; NID:g386429; PIDN:AA013938.1; PID:g4261638
 C:Comment: The GSTM1 locus for the mu isoenzyme of glutathione transferase is polymor
 stillbene oxide.
 C:Gene(s):
 A:Gene: GDB:GSTM1; GSTM16
 A:Cross-references: GDB:120020; OMIM:138350
 A:Map position: 1p13.3-1p13.3
 A:Introns: 12/3; 38/1; 59/3; 87/1; 120/3; 152/3
 C:Superfamily: glutathione transferase
 C:Keywords: dimer; polymorphism; transferase
 F:2-218/Product: glutathione transferase class mu, GSTM1 #status predicted <MAT>

Query Match	25.3%	Score 467.5;	DB 2;	Length 218;
Best Local Similarity	42.6%	Pred. No. 3.8e-31;		
Matches 92;	Conservative 39;	Mismatches 80;	Indels 5;	Gaps 1;
OY	1 MSPILGYWKIKGLVOPFRLLLEYLEEKREHLYERDEG-----DKMKNKKFELGLEFPNL 55			
DB	1 MPMLIGWMDIRGLAHAIIRLLETTDSSYEKKTYMGAPDYDSQWLNEFKLGLDFPNL 60			
OY	56 PYIIDGDKLQSMALIRYIADKHNMLGCPKERAETSMLEGAVALDIRGVSRIVASKDF 115			
DB	61 PYLIDGAHNRITQSNALICYIARKNHNLGCTEEREEKIRVDIENQTMQHMQLGMICYNPEF 120			
OY	116 ETLKVDPLSLPEMLKHFEDRLCKITLNDHTHPFMYDALDVLVLYMDPKCLDAFPK 175			
DB	121 EKLKPKYLELPKLPKLYSEFLGKRPWFAGNKTTFVDVLVLDLHRIPEPKCLDAFPN 180			
OY	176 LVCFKKRIEALPOIDKYLKSSKYIAMPLOGMQATFG 211			
DB	181 LKDFISFEGLEKISAVMKSSRFLPRPVFSKMAVWG 216			

Search completed: March 18, 2001, 05:33:19
 Job time: 2107 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2001, 05:30:37 ; Search time 71.87 seconds

(without alignments)
154,124 Million cell updates/sec

Title: US-09-402-488a-2

Perfect score: 1850
Sequence: 1 MSPILGYWKIKGLVQPTRL.....PKQSHNDGDFEIEPEEYLQ 343

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the total score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	1179	63.7	218 1	GT26_SCHUA P08515 schistosoma
2	990	53.5	218 1	GT27_SCHUA P35661 schistosoma
3	927	50.1	218 1	GT26_SCHUA P15964 schistosoma
4	710.5	38.4	220 1	GT29_FASHE P56598 fasciola he
5	699	37.8	217 1	GT27_FASHE P31670 fasciola he
6	681	36.8	217 1	GT28_FASHE P31671 fasciola he
7	680	36.8	217 1	GT26_FASHE P30112 fasciola he
8	481.5	26.0	217 1	GTM1_MOUSE P10649 mus musculu
9	481.5	26.0	217 1	GTM1_MOUSE P04905 rattus norv
10	480.5	25.8	218 1	GTM4_HUMAN P03013 homo sapien
11	477.5	25.8	217 1	GTM4_RABIT P46409 oryctolagus
12	473.5	25.6	217 1	GTM2_HUMAN P28161 homo sapien
13	473.5	25.6	217 1	GTM2_CAVPO P16413 cavia porce
14	468.5	25.3	217 1	GTM2_CRILLO P000285 cricetulus
15	468.5	25.3	217 1	GTM2_MESAU P30116 mesocricetu
16	467.5	25.3	217 1	GTM2_MOUSE P15626 mus musculu
17	465.5	25.2	217 1	GTM1_HUMAN P09488 homo sapien
18	463.5	25.1	217 1	GTM5_HUMAN P46439 homo sapien
19	461.5	24.9	217 1	GTM3_RAT P08009 rattus norv
20	455.5	24.6	217 1	GTM3_RAT P08010 rattus norv
21	452.5	24.5	217 1	GTM3_MOUSE P19639 mus musculu
22	449.5	24.3	225 1	GTM3_HUMAN P21266 homo sapien
23	437.5	23.6	224 1	GTM5_MOUSE P48774 mus musculu
24	432.5	23.4	219 1	GTM2_CHICK P20136 gallus gall
25	385.5	20.8	219 1	GTM1_DERPT P46419 dermatophag
26	367	19.8	65 1	ITH1_HIRME P01050 hirudo medi
27	359	19.4	65 1	ITHR_HIRME P28511 hirudo medi
28	355	19.2	65 1	ITHI_HIRME P28509 hirudo medi
29	353	19.1	65 1	ITHH_HIRME P28508 hirudo medi
30	352	19.0	65 1	ITHG_HIRME P28507 hirudo medi
31	352	19.0	65 1	ITHI_HIRME P28510 hirudo medi
32	349	18.9	65 1	ITHF_HIRME P28506 hirudo medi
33	346	18.7	65 1	ITHC_HIRME P28503 hirudo medi

34	344	18.6	65 1	ITHD_HIRME P28504 hirudo medi
35	339	18.3	65 1	ITHH_HIRME P28505 hirudo medi
36	329	17.8	72 1	ITH3_HIRME P09945 hirudo medi
37	311	16.8	66 1	ITH2_HIRME P09944 hirudo medi
38	300	16.2	65 1	ITHA_HIRME P28504 hirudo medi
39	251.5	13.6	208 1	GT2_DIRIM P46426 difoliaria
40	249	13.5	68 1	ITHV_HIRMA P81492 hirudinaria
41	248.5	13.4	208 1	GT2_ONCVO P46427 onchocerca
42	244	13.2	209 1	GT2_MESAU P06550 mesocricetu
43	242	13.1	207 1	GT2_PIG P80031 sus scrofa
44	240	13.0	209 1	GT2_HUMAN P09211 homo sapien
45	238	12.9	209 1	GT2_CRIMI P47954 cricetulus

ALIGNMENTS

RESULT ID	GT26_SCHUA	STANDARD:	PRT:	218 AA.
AC	P08515;			
DT	01-AUG-1988 (Rel. 08, Created)			
DT	01-AUG-1988 (Rel. 08, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	GLUTATHIONE S-TRANSFERASE 26 KDA (EC 2.5.1.18) (GST 26) (S26 ANTIGEN)			
DE	(GST CLASS-ALPHA).			
OS	Schistosoma japonicum (Blood fluke).			
OC	Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea;			
OC	Strigoidae; Schistosomatidae; Schistosomatidae; Schistosoma.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=87041520; PubMed=3095841;			
RA	Smith D.B., Davern K.M., Board P.G., Tiu W.U., Garcia E.G.,			
RA	Mitchell G.F.;			
RA	Proc. Natl. Acad. Sci. U.S.A. 84:6541-6541(1987).			
RT	[3]			
RP	X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).			
RX	MEDLINE=95276631; PubMed=7538846;			
RA	Lim K., Ho J.X., Keeling K., Gilliland G.L., Ji X., Ruker F.,			
RA	Carter D.C.;			
RT	"Three-dimensional structure of Schistosoma japonicum glutathione S-			
RT	transferase fused with a six-amino acid conserved neutralizing			
RT	epitope of gp41 from HIV."			
RL	Protein Sci. 3:2233-2244(1994).			
RN	[4]			
RP	X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).			
RX	MEDLINE=95156484; PubMed=7853399;			
RA	McGill M.A., Williams D.R., Tainer J.A.;			
RT	"Crystal structures of a schistosomal drug and vaccine target:			
RT	glutathione S-transferase from Schistosoma japonica and its complex			
RT	with the leading antischistosomal drug praziquantel."			
RL	J. Mol. Biol. 246:21-27(1995).			
CC	- FUNCTION: CONUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER			
CC	OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.			
CC	- FUNCTION: GST ISOENZYMES APPEAR TO PLAY A CENTRAL ROLE IN THE			
CC	PARASITE DETOXIFICATION SYSTEM. OTHER FUNCTIONS ARE ALSO			
CC	SUSPECTED INCLUDING A ROLE IN INCREASING THE SOLUBILITY OF			
CC	HAEMATIN IN THE PARASITE GUT.			
CC	- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GLUTATHIONE.			
CC	- SUBUNIT: HOMODIMER.			
CC	- MISCELLANEOUS: THERE ARE AT LEAST TWO ISOENZYMES OF GST IN			
CC	S. JAPONICUM.			
CC	- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. ALPHA FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL: M14654; AAB55203.1; -
CC PIR: A26484; A26484.
CC PDB: IGNE; 30-NOV-54.
CC PDB: IGTA; 07-FEB-95.
CC PDB: IGTA; 01-DEC-95.
CC PDB: 1B8X; 12-APR-99.
CC INTERPRO: IPR000521; -
CC PIR: PF00043; GST; 1.
CC Transferase: Antigen; Multigene family; 3D-structure.
CC KW SEQUENCE 218 AA; 25498 MW; 5E2AC418BD0EF13F CRC64;
SQ

Query Match 63.7%; Score 1179; DB 1; Length 218;
Best Local Similarity 100.0%; Pred. No. 3.8e-90;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSPILGYMKIKGLVPTLLLEYLEKYEHLIERDEGDKMKNKFEGLFEPNLPYYID 60
DB 1 MSPILGYMKIKGLVPTLLLEYLEKYEHLIERDEGDKMKNKFEGLFEPNLPYYID 60
QY 61 GPKVLTQSMALIRYIADKHNMLGCGPKERAETSMLEGAVIDIRYGVSRIVAKDEFTLV 120
DB 61 GPKVLTQSMALIRYIADKHNMLGCGPKERAETSMLEGAVIDIRYGVSRIVAKDEFTLV 120
QY 121 DFLSKLPKMKFEEDRLCHKTYLNGDHYTHDPFMYDALDVLVYMDPCLDAFPKLVCFK 180
DB 121 DFLSKLPKMKFEEDRLCHKTYLNGDHYTHDPFMYDALDVLVYMDPCLDAFPKLVCFK 180
QY 181 KRIEAIPOIDKTKLSSKTYIAMPLOGMOTFFGGDHPK 218
DB 181 KRIEAIPOIDKTKLSSKTYIAMPLOGMOTFFGGDHPK 218

RESULT: 2
GT27_SCHMA STANDARD: PRT: 218 AA.
AC P35661.

DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE GLUTATHIONE S-TRANSFERASE 26 KDA (EC 2.5.1.18) (GST 26) (SM26/2
DE ANTIGEN) (GST CLASS-ALPHA).
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea;
OC Strigoida; Schistosomatidae; Schistosomatidae; Schistosoma.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92131046; PubMed=1775156;
RA Wright M.D., Harrison R.A., Meider A.M., Newport G.R., Mitchell G.F.;
RT "Another 26-kilodalton glutathione S-transferase of Schistosoma
mansoni".
RL Mol. Biochem. Parasitol. 49:177-179(1991).
CC -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC -1- FUNCTION: GST ISOENZYMES APPEAR TO PLAY A CENTRAL ROLE IN THE
CC PARASITE DETOXIFICATION SYSTEM. OTHER FUNCTIONS ARE ALSO
CC SUSPECTED INCLUDING A ROLE IN INCREASING THE SOLUBILITY OF
CC HAEMATIN IN THE PARASITE GUT.
CC -1- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GLUTATHIONE.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. ALPHA FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL: M73624; -; NOT_ANNOTATED_CDS.
CC PIR: A45556; A45556.
CC HSSP: P08515; 1GTB.
CC INTERPRO: IPR000521; -
CC PIR: PF00043; GST; 1.
CC Transferase: Antigen; Multigene family.
CC KW SEQUENCE 218 AA; 25411 MW; DBD3FE9028B36185 CRC64;
SQ

Query Match 53.5%; Score 990; DB 1; Length 218;
Best Local Similarity 82.6%; Pred. No. 1.4e-74;
Matches 180; Conservative 19; Mismatches 19; Indels 0; Gaps 0;

QY 1 MSPILGYMKIKGLVPTLLLEYLEKYEHLIERDEGDKMKNKFEGLFEPNLPYYID 60
DB 1 MSPILGYMKIKGLVPTLLLEYLEKYEHLIERDEGDKMKNKFEGLFEPNLPYYID 60
QY 61 GPKVLTQSMALIRYIADKHNMLGCGPKERAETSMLEGAVIDIRYGVSRIVAKDEFTLV 120
DB 61 GPKVLTQSMALIRYIADKHNMLGCGPKERAETSMLEGAVIDIRYGVSRIVAKDEFTLV 120
QY 121 DFLSKLPKMKFEEDRLCHKTYLNGDHYTHDPFMYDALDVLVYMDPCLDAFPKLVCFK 180
DB 121 DFLSKLPKMKFEEDRLCHKTYLNGDHYTHDPFMYDALDVLVYMDPCLDAFPKLVCFK 180
QY 181 KRIEAIPOIDKTKLSSKTYIAMPLOGMOTFFGGDHPK 218
DB 181 KRIEAIPOIDKTKLSSKTYIAMPLOGMOTFFGGDHPK 218

RESULT: 3
GT26_SCHMA STANDARD: PRT: 218 AA.
AC P15964.

DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE GLUTATHIONE S-TRANSFERASE 26 KDA (EC 2.5.1.18) (GST 26) (SM26/1
DE ANTIGEN) (GST CLASS-ALPHA).
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea;
OC Strigoida; Schistosomatidae; Schistosomatidae; Schistosoma.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=PUERTO RICAN.
RX MEDLINE=90348716; PubMed=2385266;
RA Trottein F., Kieny M.P., Verwaerde C., Torpier G., Pierce R.J.,
RA Balloul J.-M., Schmitt D., Lecocq J.-P., Capron A.;
RT "Molecular cloning and tissue distribution of a 26-kilodalton
Schistosoma mansoni glutathione S-transferase".
RL Mol. Biochem. Parasitol. 41:35-44(1990).
CC -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC -1- FUNCTION: GST ISOENZYMES APPEAR TO PLAY A CENTRAL ROLE IN THE
CC PARASITE DETOXIFICATION SYSTEM. OTHER FUNCTIONS ARE ALSO
CC SUSPECTED INCLUDING A ROLE IN INCREASING THE SOLUBILITY OF
CC HAEMATIN IN THE PARASITE GUT.
CC -1- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GLUTATHIONE.
CC -1- SUBUNIT: HOMODIMER.
CC -1- TISSUE SPECIFICITY: TEGUMENT AND IN SUBTEGUMENTARY PARENCHYMAL
CC CELLS. GST 26 MAY BE ACTIVELY EXCRETED BY ADULT WORMS.

CC -1- MISCELLANEOUS: THERE ARE AT LEAST THREE ISOENZYMES OF GST IN
CC S. MANSONI.
CC -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY, ALPHA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M31106; AAA29688.1; -
CC EMBL: M26913; AAA29689.1; -
CC PIR: A45523; A45523.
CC HSSP: P08515; 1GTH.
CC INTERPRO: IPR000521; -
CC PFM: PFO0043; GST; 1.
CC Transferrase; Antigen; Multigene family.
CC SEQUENCE 218 AA; 25401 MW; 061A6548A842D6E8 CRC64;
KM
SQ

Query Match 50.1%; Score 927; DB 1; Length 218;
Best Local Similarity 79.8%; Pred. No. 2,1e-69;
Matches 174; Conservative 15; Mismatches 29; Indels 0; Gaps 0;

OY 1 MSPILGWKIKGLVPTRLLEYLEEKYEHLERDEGDMRNKKFELGLEPPLPYIID 60
DB 1 MAPKCYWKKVLCVQPTRLLEHLEETTYERAYDRNEIDAMNDKFKLGEFPNLPYIID 60
OY 61 GGVKLTQSMALIRYIADKHNMLGCPKRAEISMLEGAVLDIRGVSRAYSDEFTLVK 120
DB 61 GGVKLTQSMALIRYIADKHNMLGCPKRAEISMLEGAVLDIRGVSRAYSDEFTLVK 120
OY 121 DELSLPEMLKMFEDRLCKTYLNGDHYTHPDMFLYDALDYVLYMDPGLDAPKLYCKR 180
DB 121 DELSLPEMLKMFEDRLCKTYLNGDHYTHPDMFLYDALDYVLYMDPGLDAPKLYCKR 180
OY 121 DELSLPEMLKMFEDRLCKTYLNGDHYTHPDMFLYDALDYVLYMDPGLDAPKLYCKR 180
DB 121 DELSLPEMLKMFEDRLCKTYLNGDHYTHPDMFLYDALDYVLYMDPGLDAPKLYCKR 180
OY 181 KRIFAIPOIDKYLKSSKTYAMPLOGMOATFGGDDHPKSD 218
DB 181 KRIFAIPOIDKYLKSSKTYAMPLOGMOATFGGDDHPKSD 218
OY 181 KRIFAIPOIDKYLKSSKTYAMPLOGMOATFGGDDHPKSD 218
DB 181 KRIFAIPOIDKYLKSSKTYAMPLOGMOATFGGDDHPKSD 218

RESULT 4
GT29_FASHE STANDARD: PRT; 220 AA.
AC P56598;
ID 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GLUTATHIONE S-TRANSFERASE 26 KDA 1 (EC 2.5.1.18) (GST1) (FHL) (GST
DE CLASS-ALPHA).
OS Fasciola hepatica (Liver fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea;
OC Echinostomida; Echinostomata; Fascioloidae; Fasciolidae; Fasciola.
RN [1]
RP MEDLINE=92155306; PubMed=1740183;
RX Panaccio M., Wilson L.R., Crameri S.L., Wjiffels G.L., Spithill T.W.;
RA "Molecular characterization of cDNA sequences encoding glutathione S-
RT transferases of Fasciola hepatica.";
RN Exp. Parasitol. 74:232-237(1992).
RL [2]
RP MEDLINE=94039664; PubMed=8224094;
RX Panaccio M., Wilson L.R., Crameri S.L., Wjiffels G.L., Spithill T.W.;
RA Exp. Parasitol. 77:385-385(1993).
RN [3]
RP SEQUENCE OF 22-220 FROM N.A.
RA Crameri S.;
RP Patent number WO9008819, 09-AUG-1990.
CC -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.

CC -1- FUNCTION: GST ISOENZYMES APPEAR TO PLAY A CENTRAL ROLE IN THE
CC PARASITE DETOXIFICATION SYSTEM. OTHER FUNCTIONS ARE ALSO
CC SUSPECTED INCLUDING A ROLE IN INCREASING THE SOLUBILITY OF
CC HAEMATIN IN THE PARASITE GUT.
CC -1- CATALYTIC ACTIVITY: RX + GLUTATHIONE -> HX + R-S-GUTATHIONE.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY, ALPHA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: A00993; CA00118.1; -
CC INTERPRO: IPR000521; -
CC PFM: PFO0043; GST; 1.
CC Transferrase; Antigen; Multigene family.
CC SEQUENCE 220 AA; 25598 MW; 27B9F150B75D101F CRC64;
KM
SQ

Query Match 38.4%; Score 710.5; DB 1; Length 220;
Best Local Similarity 59.7%; Pred. No. 1.4e-51;
Matches 129; Conservative 32; Mismatches 54; Indels 1; Gaps 1;

OY 5 LGVWKIKGLVPTRLLEYLEEKYEHLERDEGDMRNKKFELGLEPPLPYIIDGVK 64
DB 4 LGVWKIKGLVPTRLLEYLEEKYEHLERDEGDMRNKKFELGLEPPLPYIIDGVK 64
OY 65 LTQSMALIRYIADKHNMLGCPKRAEISMLEGAVLDIRGVSRAYSDEFTLVK 124
DB 65 LTQSMALIRYIADKHNMLGCPKRAEISMLEGAVLDIRGVSRAYSDEFTLVK 124
OY 125 KLPEMLKMFEDRLCKTYLNGDHYTHPDMFLYDALDYVLYMDPGLDAPKLYCKR 184
DB 125 KLPEMLKMFEDRLCKTYLNGDHYTHPDMFLYDALDYVLYMDPGLDAPKLYCKR 184
OY 125 KLPEMLKMFEDRLCKTYLNGDHYTHPDMFLYDALDYVLYMDPGLDAPKLYCKR 184
DB 125 KLPEMLKMFEDRLCKTYLNGDHYTHPDMFLYDALDYVLYMDPGLDAPKLYCKR 184
OY 185 ALPQIDKYLKSSKTYAMPLOGMOATFGGDDHPKSD 220
DB 185 ALPQIDKYLKSSKTYAMPLOGMOATFGGDDHPKSD 220
OY 185 ALPQIDKYLKSSKTYAMPLOGMOATFGGDDHPKSD 220
DB 185 ALPQIDKYLKSSKTYAMPLOGMOATFGGDDHPKSD 220

RESULT 5
GT27_FASHE STANDARD: PRT; 217 AA.
AC P31670;
ID 01-JUL-1993 (Rel. 26, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GLUTATHIONE S-TRANSFERASE 26 KDA 47 (EC 2.5.1.18) (GST47) (FH47)
DE (GST CLASS-ALPHA).
OS Fasciola hepatica (Liver fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea;
OC Echinostomida; Echinostomata; Fascioloidae; Fasciolidae; Fasciola.
RN [1]
RP MEDLINE=92155306; PubMed=1740183;
RX Panaccio M., Wilson L.R., Crameri S.L., Wjiffels G.L., Spithill T.W.;
RA "Molecular characterization of cDNA sequences encoding glutathione S-
RT transferases of Fasciola hepatica.";
RN Exp. Parasitol. 74:232-237(1992).
RL [2]
RP MEDLINE=94039664; PubMed=8224094;
RX Panaccio M., Wilson L.R., Crameri S.L., Wjiffels G.L., Spithill T.W.;
RA Exp. Parasitol. 77:385-385(1993).
RN [3]
RP SEQUENCE OF 22-220 FROM N.A.
RA Crameri S.;
RP Patent number WO9008819, 09-AUG-1990.
CC -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.


```

AC P30112:
DT 01-APR-1993 (rel. 25, Created)
DT 15-DEC-1998 (rel. 37, Last sequence update)
DT 15-DEC-1998 (rel. 37, Last annotation update)
DE GLUTATHIONE S-TRANSFERASE 26 KDA 51 (EC 2.5.1.18) (GSM51) (FHS1)
DE (GST CLASS-ALPHA).
OS Fasciola hepatica (Liver fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea;
OC Echinostomida; Echinostomata; Fascioloidae; Fasciolidae; Fasciola.
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=92155306; PubMed=1740183;
RA Panaccio M., Wilson L.R., Cramer S.L., Wiffels G.L., Spithill T.W.;
RT "Molecular characterization of cDNA sequences encoding glutathione S-
RT transferases of Fasciola hepatica."
RL Exp. Parasitol. 74:232-237(1992).
RN 12
RP ERRATUM.
RX MEDLINE=94039664; PubMed=8224094;
RA Panaccio M., Wilson L.R., Cramer S.L., Wiffels G.L., Spithill T.W.;
RL Exp. Parasitol. 77:385-385(1993).
RN 13
RP SEQUENCE OF 9-217 FROM N.A.
RX MEDLINE=93228188; PubMed=7682383;
RA Muro A., Rodriguez-Molina J.R., Hillyer G.V.;
RT "Sequence analysis of a Fasciola hepatica glutathione S-transferase
RT cDNA clone."
RL Am. J. Trop. Med. Hyg. 48:457-463(1993).
CC -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC -1- FUNCTION: GST ISOENZYMES APPEAR TO PLAY A CENTRAL ROLE IN THE
CC PARASITE DETOXIFICATION SYSTEM. OTHER FUNCTIONS ARE ALSO
CC SUSPECTED INCLUDING A ROLE IN INCREASING THE SOLUBILITY OF
CC HAEMATIN IN THE PARASITE GUT.
CC -1- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GUTATHIONE.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. ALPHA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M77682; AAA29141.1;
DR HSSP: P08515; IGNE.
DR INTERPRO: IPR000521;
DR PFAM: PF00043; GST_1.
KW Transferase; Antigen; Multigene family.
FT INIT MET 0 BY SIMILARITY.
FT CONFLICT 83 83 T->S (IN REF. 3).
FT CONFLICT 134 134 N->D (IN REF. 3).
FT CONFLICT 146 146 P->T (IN REF. 3).
SQ SEQUENCE 217 AA; 25242 MW; EFCABFF7454EDC26 CRC64;

```

Query Match 36.8%; Score 680; DB 1; Length 217;
 Best Local Similarity 58.7%; Pred. No. 4,4e-49;
 Matches 125; Conservative 32; Mismatches 56; Indels 0; Gaps 0;

```

DB 124 ELPTLKMSNFIQDNRHYLGSPVSHVDFMVAELDCIRLAQCLEDPFKLEFKSRIE 183
OY 185 AIPQIDKYKSSKYIAMPLOQWQATFGGDDHP 217
DB 184 DLPIKAYMESEKFIKWPLNSWIASFCGGDAAP 216

RESULT 8
ID GTM1_MOUSE STANDARD; PRT; 217 AA.
AC P10649;
DT 01-JUL-1989 (rel. 11, Created)
DT 01-JUL-1989 (rel. 11, Last sequence update)
DT 15-DEC-1998 (rel. 37, Last annotation update)
DE GLUTATHIONE S-TRANSFERASE GTR.7 (EC 2.5.1.18) (GST 1-1) (GST CLASS-
DE MU).
GN GSTM1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=88330838; PubMed=3417659;
RA Pearson W.R., Reinhart J., Sisk S.C., Anderson K.S., Adler P.N.;
RT "Tissue-specific induction of murine glutathione transferase mRNAs by
RT butylated hydroxyanisole."
RL J. Biol. Chem. 263:13324-13332(1988).
RN 12
RP SEQUENCE FROM N.A.
RX MEDLINE=90094327; PubMed=2689439;
RA Townsend A.J., Goldsmith M.E., Pickett C.B., Cowan K.H.;
RT "Isolation, characterization, and expression in Escherichia coli of
RT two murine Mu class glutathione S-transferase cDNAs homologous to the
RT rat subunits 3 (Yb1) and 4 (Yb2)."
RL J. Biol. Chem. 264:21582-21590(1989).
RN 13
RP SEQUENCE FROM N.A.
RX MEDLINE=93290350; PubMed=8512323;
RA Reinhardt J., Pearson W.R.;
RT "The structure of two murine class-mu glutathione transferase genes
RT coordinately induced by butylated hydroxyanisole."
RL Arch. Biochem. Biophys. 303:383-393(1993).
RN 14
RP PRELIMINARY SEQUENCE OF 1-40.
RX MEDLINE=83109018; PubMed=6822548;
RA Pearson W.R., Windle J.J., Morrow J.F., Benson A.M., Talalay P.;
RT "Increased synthesis of glutathione S-transferases in response to
RT anticarcinogenic antioxidants. Cloning and measurement of messenger
RT RNA."
RL J. Biol. Chem. 258:2052-2062(1983).
RN 15
RP PRELIMINARY SEQUENCE OF 1-24.
RX MEDLINE=86042634; PubMed=3864155;
RA Mannervik B., Alin P., Guthenberg C., Jonsson H., Tahir M.K.,
RA Warholm M., Joernvall H.;
RT "Identification of three classes of cytosolic glutathione transferase
RT common to several mammalian species: correlation between structural
RT data and enzymatic properties."
RL Proc. Natl. Acad. Sci. U.S.A. 82:7202-7206(1985).
RN 16
RP CHARACTERIZATION.
RX STRAIN-CD-1; TISSUE=LIVER;
RX MEDLINE=96189427; PubMed=8605288;
RA Mitchell A.E., Morin D., Lane M.W., Jones A.D.;
RT "Purification, mass spectrometric characterization, and covalent
RT modification of murine glutathione S-transferases."
RL Chem. Res. Toxicol. 8:1054-1062(1995).
CC -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC -1- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GUTATHIONE.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- MASS SPECTROMETRY: MW=25838.4; MW_ERR=2; METHOD=-ELECTROSPRAY.

```



```
CC ----- This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercialisation  
CC entities requires a license agreement (see http://www.isb-slb.ch/announce/  
CC or send an email to license@isb-slb.ch).  
CC -----  
DR EMBL; X04229; CAA27811.1; -;  
DR EMBL; M11719; AAA41287.1; -;  
DR EMBL; J02810; AAA41293.1; -;  
DR PIR; A24085; A24085.  
DR PIR; A25510; A25510.  
DR PIR; A29794; A29794.  
DR PIR; S17167; S17167.  
DR PDB; 1G5B; 3I-OCT-93.  
DR PDB; 1G5C; 3I-OCT-93.  
DR PDB; 2GST; 3I-OCT-93.  
DR PDB; 3GST; 3I-JAN-94.  
DR PDB; 4GST; 3I-OCT-93.  
DR PDB; 5GST; 3I-OCT-93.  
DR PDB; 6GSU; 08-NOV-96.  
DR PDB; 6GSU; 08-NOV-96.  
DR PDB; 6GSU; 08-NOV-96.  
DR PDB; 6GSW; 08-NOV-96.  
DR PDB; 6GSX; 08-NOV-96.  
DR PDB; 6GSY; 08-NOV-96.  
DR PDB; 5FWG; 27-JAN-99.  
DR INTERPRO; IPR000521; -;  
DR INTERPRO; IPR003081; -;  
DR PFAM; PF00043; GST; 1.  
DR PRINTS; PRO1267; GSTRNSFRASEM.  
KW Transferrase; Multigene family; 3D-structure; Olfaction.  
FT INIT MET 0 0  
FT MUTAGEN 86 86 C->S. NO CHANGE IN ACTIVITY.  
FT CONFLICT 168 168 I-> N (IN REF. 3).  
FT CONFLICT 198 199 KS -> NC (IN REF. 2).  
FT STRAND 2 7  
FT TURN 11 13  
FT HELIX 14 22  
FT TURN 23 24  
FT STRAND 27 32  
FT TURN 37 39  
FT HELIX 43 46  
FT TURN 47 50  
FT STRAND 61 64  
FT TURN 65 66  
FT STRAND 67 70  
FT HELIX 72 82  
FT TURN 83 84  
FT HELIX 90 114  
FT TURN 115 115  
FT TURN 117 118  
FT HELIX 119 128  
FT TURN 129 129  
FT HELIX 130 141  
FT TURN 142 143  
FT TURN 154 154  
FT HELIX 155 169  
FT TURN 171 176  
FT HELIX 178 188  
FT TURN 189 189  
FT HELIX 191 196  
FT TURN 197 198  
FT TURN 200 201  
FT TURN 210 211  
SQ SEQUENCE 217 AA; 25782 MW; 2ACE8D49DA785118 CRC64;
```

```

OY 4 ILGVKIKGLVQPTPLLEYLEEKKEEHLBERDEG-----DKMRNKKPEELGLEPNNLPY 58
DB 3 ILGVNWNGLTHPIPLLELTETDTSYEERKYAMGDAPDYDRSOWLEKFKLGLDPENNLPYL 62
OY 59 IDGDVKLTQSMNAIITFYIADKHNMLGCGCKPERAETSMLEGCAYDLIRYGVSRVASKDFEPL 118
DB 63 IDGSKRIITQSMALIMRYLARKHHLGCETEEERIRADIVENQWMDNNMOLIMLCYNDFEKO 122
OY 119 KYDELSTKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVLVYMDPCLDAEPFLVC 178
DB 123 KPEEFKLTPEKKMKLTSEFLGKRPFAGKGVYVDFLAVDILDQYHIFEPCKDAPENLKD 182
OY 179 FKRIEAIPOIDKYLVSSKSYIAMP 203
DB 183 FLAREEGLKISAVWKSRSRYLSTPI 207

RESULT 10
GTW4_HUMAN
ID GTW4_HUMAN STANDARD PRT: 218 AA.
AC Q03013.1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE GLUTATHIONE S-TRANSFERASE MU 4 (EC 2.5.1.18) (GSTW4-4) (GTS-MU2)
DE (GST CLASS-MU).
GN GSTW4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RN [1]
RP SEQUENCE FROM N.A..
RX MEDLINE=93352467; PubMed=8349586;
RA Comstock K.E., Johnson K.J., Rifenbery D., Henner W.D.;
RT "Isolation and analysis of the gene and cDNA for a human Mu class
RT glutathione S-transferase, GSTW4.";
RL J. Biol. Chem. 268:16958-16965(1993).
RN [2]
RP SEQUENCE FROM N.A..
RX MEDLINE=93228631; PubMed=8471052;
RA Zhong S., Spurr N.K., Hayes J.D., Wolf C.R.;
RT "Deduced amino acid sequence, gene structure and chromosomal location
RT of a novel human class Mu glutathione S-transferase, GSTW4.";
RL Biochem. J. 291:41-50(1993).
RN [3]
RP SEQUENCE OF 39-160 FROM N.A..
RX TISSUE=Lymphocytes;
RC MEDLINE=91174774; PubMed=2006920;
RA Taylor J.B., Oliver J., Sherrington R., Pemble S.E.;
RT "Structure of human glutathione S-transferase class Mu genes.";
RL Biochem. J. 274:587-593(1991).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=94263230; PubMed=8203914;
RA Comstock K.E., Widensten M., Hao X.Y., Henner W.D., Manerovic B.;
RT "A comparison of the enzymatic and physicochemical properties of
RT human glutathione transferase M4-4 and three other human Mu class
RT enzymes.";
RL Arch. Biochem. Biophys. 311:487-495(1994).
CC -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC ACTIVE ON 1-CHLORO-2,4-DINITROBENZENE.
CC -1- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GUTATHIONE.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN A WIDE VARIETY OF TISSUES.
CC -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```


modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M96234; AAA57347.1; -
DR EMBL; M96233; AAA57346.1; -
DR EMBL; X68637; CAA48637.1; -
DR EMBL; X56837; CAA40167.1; -
DR HSSP; P28161; 1HNH.
DR MIM; J38333; -
DR INTERPRO; IPR000521; -
DR INTERPRO; IPR003081; -
DR PFAM; PF00043; GST; 1.
DR PRINTS; PR01267; GSTRNSFRASEM.
KW Transferase; Multigene family.
FT CONFLICT 2 SMT -> PMI (IN REF. 2).
FT CONFLICT 17 I -> M (IN AAA57346).
FT CONFLICT 37 D -> G (IN REF. 2).
FT CONFLICT 212 V -> M (IN REF. 2).
SO SEQUENCE 218 AA; 25561 MW; 35E96FA54D566B1E CRC64;

Query Match 26.0%; Score 480.5; DB 1; Length 218;
Best Local Similarity 44.0%; Pred. No. 1.2e-32;
Matches 95; Conservative 37; Mismatches 79; Indels 5; Gaps 1;

OY 1 MSPILGYKIKGLVPTRLLEYLEEKEEHLERDEG-----DKRRKKFELGLEPPLYYI 55
DB 1 MSMTGLYWDIRGLAHAIIRLLELYTDTSYEEKRYMGDAPNDOSKWSKEFTLGLDFPL 60
OY 56 PYIDGDKLTQSMATIRIYADKHNMLGCPKERAISMLEGAVDIRGVSRIVASKDF 115
DB 61 PYLIDGAKRITDSNAILCYIARKHNLGCTEEBEKIRVDILEQANDVSQGLARVCSPDF 120
OY 116 ETLKVDLSKIDPEMKFEDRLCHKTYLNGDHYTHPDMFLYDALDVLVYMDPCIDAPFK 175
DB 121 EKLKPEYLEELPTMQHESQFLGKRPWEVGDKITFVDLAVDLDLHRIEFENCLDAEPN 180
OY 176 LVCFKKRIEALPDIDKYLKSSKYIAMPLOGQATFG 211
DB 181 LKDFISREGLEKISAYMKSSRFELPKPLRYAVAWG 216

RESULT 11
GTMU_RABIT STANDARD; PRT: 217 AA.
AC P46409;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 33, Last annotation update)
DE GLUTATHIONE S-TRANSFERASE MU 1 (EC 2.5.1.18) (GST MU 1) (GST CLASS-MU).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=95251394; PubMed=7733673;
RA Lee S.H., Lee S.H., Han J.S., Kim Y.S., Koh J.K.;
RT "Cloning and expression of a cDNA for mu-class glutathione
S-transferase from rabbit liver."
RA Arch. Biochem. Biophys. 318:424-429(1995).
RT Arch. Biochem. Biophys. 318:424-429(1995).
CC - FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC - CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GUTATHIONE.
CC - SUBUNIT: HOMODIMER.
CC - SUBCELLULAR LOCATION: CYTOPLASMIC.
CC - TISSUE SPECIFICITY: WELL EXPRESSED IN RABBIT LIVER, BRAIN, AND
KIDNEY.
CC - SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
CC - TISSUE SPECIFICITY: THIS IS A MUSCLE ISOZYME.
CC - SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; I23766; AAA69665.1; -
DR HSSP; P28161; 1HNH.
DR INTERPRO; IPR000521; -
DR INTERPRO; IPR003081; -
DR PFAM; PF00043; GST; 1.
DR PRINTS; PR01267; GSTRNSFRASEM.
KW Transferase; Multigene family.
FT INIT MET 0 BY SIMILARITY.
SO SEQUENCE 217 AA; 25286 MW; 785EA932C543ECB2 CRC64;

Query Match 25.8%; Score 477.5; DB 1; Length 217;
Best Local Similarity 44.6%; Pred. No. 2.1e-32;
Matches 95; Conservative 36; Mismatches 77; Indels 5; Gaps 1;

OY 5 LGYKIKGLVPTRLLEYLEEKEEHLERDEG-----DKRRKKFELGLEPPLYYI 59
DB 4 LGYWDVRLALPIRLMLLELYTDTSYEEKRYMGDAPNDOSKWSKEFTLGLDFPL 63
OY 60 DGDVLTQSMATIRIYADKHNMLGCPKERAISMLEGAVDIRGVSRIVASKDF 119
DB 64 DGTAKRLTQSMATIRIYADKHNMLGCPKERAISMLEGAVDIRGVSRIVASKDF 123
OY 120 VDLKSLPEMKFEDRLCHKTYLNGDHYTHPDMFLYDALDVLVYMDPCIDAPFK 179
DB 124 PEYLKGLPEKQLVYSGFLSPWFAKDKITFADFLVYDVLQNRIFVGCIDAPFN 183
OY 180 KKRIPALPDIDKYLKSSKYIAMPLOGQATFG 212
DB 184 HVREGLPKISAYMKSSRFELPKPLRYAVAWG 216

RESULT 12
GTW2_HUMAN STANDARD; PRT: 217 AA.
AC P28161;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GLUTATHIONE S-TRANSFERASE MU 2 (EC 2.5.1.18) (GSTW2-2) (GST CLASS-MU).
GN GSTW2 OR GST4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Carnivora; Homiidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=91239584; PubMed=2034681;
RX Vorachek W.R., Pearson W.R., Rule G.S.;
RT "Cloning, expression, and characterization of a class-mu glutathione
S-transferase from human muscle, the product of the GST4 locus."
RA Proc. Natl. Acad. Sci. U.S.A. 88:4443-4447(1991).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).
RX MEDLINE=94238693; PubMed=8182750;
RA Raghunathan S., Chandross R.J., Kretsinger R.H., Allison T.J.,
RA Penington C.J., Rule G.S.;
RT "Crystal structure of human class mu glutathione transferase GSTW2-2.
Effects of lattice packing on conformational heterogeneity."
RT J. Mol. Biol. 238:815-832(1994).
CC - FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC - CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GUTATHIONE.
CC - SUBUNIT: HOMODIMER.
CC - SUBCELLULAR LOCATION: CYTOPLASMIC.
CC - TISSUE SPECIFICITY: THIS IS A MUSCLE ISOZYME.
CC - SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
CC - SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.


```

OS Cricetus longicaudatus (Long-tailed hamster) (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
RN [1]
RP SEQUENCE FROM N.A. PubMed=2226437;
RX MEDLINE=91031445; Hyrien O., Debatisse M., Buttin G.;
RA de Saint Vincent B.R., Hyrien O., glutathione S-transferase genes and an
RT "amplification of mu class glutathione S-transferase gene in coformycin-resistant Chinese hamster
RT fibroblasts."
RL Eur. J. Biochem. 193;19-24(1990).
CC -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC -1- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GUTATHIONE.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X57489; CAA40726.1; -.
CC PIR: S13202; S13202.
CC HSSP: P04905; 6GST.
CC INTERPRO: IPR000521; -.
CC INTERPRO: IPR003081; -.
CC PFM: PF00043; GST; 1.
CC PRINTS: PR01267; GSTINSTRASEM.
CC TRANSFERASE; Multigene family.
CC INIT MET 0
CC SEQUENCE 217 AA; 25688 MW; 34D8A8FBDD8627C0 CRC64;

Query Match 25.3%; Score 468.5; DB 1; Length 217;
Best Local Similarity 43.4%; Pred. No. 1.1e-31;
Matches 89; Conservative 36; Mismatches 75; Indels 5; Gaps 1;

4 ILGVWIKIGVOPTRLLLEYLEEYEHLYERDEG-----DKRNKKFELGLEPMLPYI 58
Db 3 ILGVWYKIGVOPTRLLLEYLEEYEHLYERDEG-----DKRNKKFELGLEPMLPYI 62
QY 59 IDGVKLTOSMATIRYADKHNMLGGCPKRAEISMEGAVIDIRYGVSRIVAKDFETL 118
Db 63 IDGSHKITQSNALIRYARKNHLCGETEERIRIVDIVENAMOTRMOLIMLCVNPDEKO 122
QY 119 KVDFSLKPEMLKMFEDRCHKTYLNGDHTHPDMLYDALDVLLVMDPMDCAFRLVCF 178
Db 123 KPEFLKTIPIKRMVSEFLGRKRWFPADKVTLLCGFLAYDVLQYQEFKCLDPFRLND 182
QY 179 FKRIEAIPOIDKYLKSSKYIAMPL 203
Db 183 FLAFEGLKRTISAVMKTSRLRPI 207

RESULT 15
GTWU_MESAU STANDARD; PRT; 217 AA.
ID GTWU_MESAU STANDARD; PRT; 217 AA.
AC P30116;
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1998 (Rel. 35, Last annotation update)
DE GLUTATHIONE S-TRANSFERASE (EC 2.5.1.18) (GST CLASS-MU).
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=SMOOTH MUSCLE;
RX MEDLINE=92335246; PubMed=1631097;
RA Fan W.M., Trifiletti R., Norris J.S., Cooper T.M.;
RT "Cloning of a mu-class glutathione S-transferase gene and
RT identification of the glucocorticoid regulatory domains in its 5'
RT flanking sequence."
RL Proc. Natl. Acad. Sci. U.S.A. 89;6104-6108(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=SMOOTH MUSCLE;
RX MEDLINE=92049380; PubMed=1944302;
RA Norris J.S., Schwartz D.A., MacLeod S.L., Fan W.M., O'Brien T.J.,
RA Harris S.E., Trifiletti R., Cornett L.E., Cooper T.M., Levi W.M.,
RA Smith R.G.;
RT "Cloning of a mu-class glutathione S-transferase complementary DNA
RT and characterization of its glucocorticoid inducibility in a smooth
RT muscle tumor cell line."
RL Mol. Endocrinol. 5;979-986(1991).
CC -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC -1- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-GUTATHIONE.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M59773; AAA37075.1; -.
CC PIR: A23732; A23732.
CC HSSP: P04905; 6GST.
CC INTERPRO: IPR000521; -.
CC INTERPRO: IPR003081; -.
CC PFM: PF00043; GST; 1.
CC PRINTS: PR01267; GSTINSTRASEM.
CC TRANSFERASE; Multigene family.
CC INIT MET 0
CC SEQUENCE 217 AA; 25558 MW; C59812AA4A30B6375 CRC64;

Query Match 25.3%; Score 468.5; DB 1; Length 217;
Best Local Similarity 44.7%; Pred. No. 1.1e-31;
Matches 89; Conservative 37; Mismatches 68; Indels 5; Gaps 1;

5 LGVWIKIGVOPTRLLLEYLEEYEHLYERDEG-----DKRNKKFELGLEPMLPYI 59
Db 4 LGVWIRIGLAIHRIQLLETDTSYEKKYTMGDAPNFRSQWLNKRFKLGDFPMLPYI 63
QY 60 DGVKLTOSMATIRYADKHNMLGGCPKRAEISMEGAVIDIRYGVSRIVAKDFETL 119
Db 64 DGSHTKTSNAILIRYARKNHLCGETEERIRIVDIVENAMOTRMOLIMLCVNPDEKO 123
QY 120 VDFSLKPEMLKMFEDRCHKTYLNGDHTHPDMLYDALDVLLVMDPMDCAFRLVCF 179
Db 124 PEYLEGLPKMLLYSEFLGRKRWFPADKVTLLCGFLAYDVLQYQEFKCLDPFRLND 183
QY 180 KRIEAIPOIDKYLKSSKY 198
Db 184 LAFEGLKRTISAVMKTSR 202

Search completed: March 18, 2001, 05:38:37
Job time: 480 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2001, 05:29:32 ; Search time 74.32 Seconds
(without alignments)
540.935 million cell updates/sec

Title: US-09-402-488a-2
Perfect score: 1650
Sequence: 1 MSPILGWKIKGLVQPTRL.....PKPSHNDGFEELPEEXLQ 343

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues
Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1244.5	67.3	245	2	009193
2	1226	66.3	262	2	009195
3	1226	66.3	268	2	009194
4	1202.5	65.0	243	2	009192
5	1202.5	65.0	244	2	009191
6	1049	56.7	218	5	094745
7	718	38.8	218	5	025595
8	698	37.7	218	5	09XYL9
9	512.5	27.2	223	6	09NMV4
10	502.5	25.9	218	11	035660
11	478.5	25.8	218	6	09TSM5
12	477.5	25.6	218	6	09TSM4
13	474.5	25.6	219	5	027653
14	473.5	25.6	219	5	016058
15	464.5	25.1	218	11	09WU21
16	464.5	24.2	225	4	060550
17	448.5	23.6	195	4	005465
18	436.5	23.5	225	11	0921B2
19	434.5	23.5	225	11	0921B2

20	407.5	22.0	188	6	09WZB4	09WZB4 capra hircu
21	377.5	20.4	129	5	090582	09u582 psoroptes o
22	288.5	15.6	125	6	029583	029583 sus scrofa
23	256.5	13.9	208	5	002636	002636 brugia mala
24	247.5	13.4	208	5	027711	027711 onchocerca
25	247	13.4	84	5	007557	007557 hirudinaria
26	246	13.3	210	13	P81942	P81942 bufo bufo
27	240	13.0	210	4	Q15690	Q15690 homo sapien
28	240	13.0	210	4	000460	000460 homo sapien
29	236	12.8	180	5	09NHB2	09nhb2 mytilus edu
30	233	12.6	209	6	09TTY8	09tty8 capra hircu
31	230	12.4	84	5	007558	007558 hirudinaria
32	213.5	11.5	216	13	Q9W647	Q9w647 oncorhynch
33	212.5	11.5	82	4	Q9UE37	Q9ue37 homo sapien
34	203	11.0	366	6	09N1P5	09n1p5 bubalus bub
35	197.5	10.7	345	6	027951	027951 bos taurus
36	192.5	10.4	208	5	P91505	P91505 caenorhabdi
37	191.5	10.4	226	11	Q90LX3	Q90lx3 rattus norv
38	189.5	10.2	221	11	P70686	P70686 mesocricetu
39	187	10.1	210	5	Q9N4X8	Q9n4x8 caenorhabdi
40	186.5	10.1	222	6	09N2J6	09n2j6 ovine aries
41	182.5	9.9	208	5	061750	061750 caenorhabdi
42	182.5	9.9	223	6	029057	029057 sus scrofa
43	180.5	9.8	199	11	Q90HE7	Q90he7 mus musculu
44	179.5	9.7	222	4	Q9NTY6	Q9nty6 homo sapien
45	179.5	9.7	222	6	Q9XS30	Q9xs30 ovine aries

ALIGNMENTS

RESULT 1
ID 009193 PRELIMINARY; PRT; 245 AA.
AC 009193;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE GLUTATHIONE S-TRANSFERASE.
GN GST.
OS Escherichia coli.
OG Plasmid pGEX-6P-3.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN 111
RP SEQUENCE FROM N.A.
RX MEDLINE=88329742; PubMed=3047011;
RA Smith D.B., Johnson K.S.;
RT "Single-step purification of polypeptides expressed in Escherichia
RT coli as fusions with glutathione S-transferase.";
RL Gene 67:31-40(1988).
RN 121
RP SEQUENCE OF 221-227 FROM N.A.
RX MEDLINE=90264387; PubMed=2160953;
RA Cordingley M.G., Callahan P.L., Sardana V.V., Garsky V.M.,
RA Colombo R.J.;
RT "Substrate requirements of human rhinovirus 3C protease for peptide
RT cleavage in vitro.";
RL J. Biol. Chem. 265:9062-9065(1990).
RN 131
RP SEQUENCE FROM N.A.
RA Dunst R.W., Bell P.A., English J.D.;
RX Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U78874; AAB37352.1; -;
DR HSSP: P08515; IGENE.
DR INTERPRO: IPR000521; -;
DR PFM: PF00043; GST_1.
KW Transferase; Plasmid.
SQ SEQUENCE 245 AA; 28327 MW; D51470B8A48C6CC2 CRC64;

Query Match

67.3%; Score 1244.5; DB 2; Length 245;

Best Local Similarity 97.1%; Pred. No. 2,1e-94;
Matches 234; Conservative 0; Mismatches 2; Indels 5; Gaps 1;

QY 1 MSPILGYKIKGLVPTLLLEYLEEKYEENLYERDEGDKMNKFFELGLEPNNLPYYID 60
DB 1 MSPILGYKIKGLVPTLLLEYLEEKYEENLYERDEGDKMNKFFELGLEPNNLPYYID 60
QY 61 GVKLTQSMATIRYIADKHNMLGGCPKERAETSMLEGAVALDIRYGVSRAYSKDDEFTLV 120
DB 61 GVKLTQSMATIRYIADKHNMLGGCPKERAETSMLEGAVALDIRYGVSRAYSKDDEFTLV 120
QY 121 DFLSKLPEMLKMFEDRLCHKTYLNGDHYTHPDPMYDALDVLVYMDPMCLDAFPKLVCFK 180
DB 121 DFLSKLPEMLKMFEDRLCHKTYLNGDHYTHPDPMYDALDVLVYMDPMCLDAFPKLVCFK 180
QY 181 KRIEAIPOIDKYLKSSKYIAMPLOGMQATFFGGDHPKSDLV-----PRGSPNSRVDSSG 235
DB 181 KRIEAIPOIDKYLKSSKYIAMPLOGMQATFFGGDHPKSDLV-----PRGSPNSRVDSSG 240
QY 236 R 236
DB 241 R 241

RESULT 2

009195 PRELIMINARY: PRT: 262 AA.

AC 009195:
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE GST-6HIS.
GN GST.
OS Escherichia coli.
OC Plasmid pGEX-2T-6His-PL2.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RA Hipskind R.H., Delaney P.;
RL Submitted (Feb-1997) to the EMBL/GenBank/DBJ databases.
RN [12]
RP SEQUENCE FROM N.A.
RA Cahill M.A.;
RL Submitted (Jan-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U84572; AAB41883.1; -
DR HSSP: P08515; IONE.
DR INTERPRO: IPR000521; -
DR PFAM: PF00043; GST; 1.
SQ SEQUENCE 262 AA; 30392 MW; BA4C4593CB20226E CRC64;

Query Match 66.3%; Score 1226; DB 2; Length 262;
Best Local Similarity 100.0%; Pred. No. 7.6e-93;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSPILGYKIKGLVPTLLLEYLEEKYEENLYERDEGDKMNKFFELGLEPNNLPYYID 60
DB 1 MSPILGYKIKGLVPTLLLEYLEEKYEENLYERDEGDKMNKFFELGLEPNNLPYYID 60
QY 61 GVKLTQSMATIRYIADKHNMLGGCPKERAETSMLEGAVALDIRYGVSRAYSKDDEFTLV 120
DB 61 GVKLTQSMATIRYIADKHNMLGGCPKERAETSMLEGAVALDIRYGVSRAYSKDDEFTLV 120
QY 121 DFLSKLPEMLKMFEDRLCHKTYLNGDHYTHPDPMYDALDVLVYMDPMCLDAFPKLVCFK 180
DB 121 DFLSKLPEMLKMFEDRLCHKTYLNGDHYTHPDPMYDALDVLVYMDPMCLDAFPKLVCFK 180
QY 181 KRIEAIPOIDKYLKSSKYIAMPLOGMQATFFGGDHPKSDLVPRGSP 227
DB 181 KRIEAIPOIDKYLKSSKYIAMPLOGMQATFFGGDHPKSDLVPRGSP 227

RESULT 3

009194 PRELIMINARY: PRT: 268 AA.

AC 009194:
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE GST-6HIS.
GN GST.
OS Escherichia coli.
OC Plasmid pGEX-2T-6H.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RA Cahill M.A.;
RL Submitted (Jan-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U84571; AAB41882.1; -
DR HSSP: P08515; IONE.
DR INTERPRO: IPR000521; -
DR PFAM: PF00043; GST; 1.
SQ SEQUENCE 268 AA; 31000 MW; 85820375FE60E625 CRC64;

Query Match 66.3%; Score 1226; DB 2; Length 268;
Best Local Similarity 100.0%; Pred. No. 7.8e-93;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSPILGYKIKGLVPTLLLEYLEEKYEENLYERDEGDKMNKFFELGLEPNNLPYYID 60
DB 1 MSPILGYKIKGLVPTLLLEYLEEKYEENLYERDEGDKMNKFFELGLEPNNLPYYID 60
QY 61 GVKLTQSMATIRYIADKHNMLGGCPKERAETSMLEGAVALDIRYGVSRAYSKDDEFTLV 120
DB 61 GVKLTQSMATIRYIADKHNMLGGCPKERAETSMLEGAVALDIRYGVSRAYSKDDEFTLV 120
QY 121 DFLSKLPEMLKMFEDRLCHKTYLNGDHYTHPDPMYDALDVLVYMDPMCLDAFPKLVCFK 180
DB 121 DFLSKLPEMLKMFEDRLCHKTYLNGDHYTHPDPMYDALDVLVYMDPMCLDAFPKLVCFK 180
QY 181 KRIEAIPOIDKYLKSSKYIAMPLOGMQATFFGGDHPKSDLVPRGSP 227
DB 181 KRIEAIPOIDKYLKSSKYIAMPLOGMQATFFGGDHPKSDLVPRGSP 227

RESULT 4

009192 PRELIMINARY: PRT: 243 AA.

AC 009192:
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE GLUTATHIONE S-TRANSFERASE.
GN GST.
OS Escherichia coli.
OC Plasmid pGEX-6P-2.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=88329742; Pubmed=3047011;
RA Smith D.B., Johnson K.S.;
RT "Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase.";
RL Gene 67:31-40(1988).
RN [12]
RP SEQUENCE OF 221-227 FROM N.A.
RX MEDLINE=90264387; Pubmed=2160953;
RA Cordingley M.C., Callahan P.L., Sardana V.V., Garsky V.M.,
RA Colonna R.J.;

RT "Substrate requirements of human rhinovirus 3C protease for peptide
cleavage in vitro.";
RL J. Biol. Chem. 265:9062-9065(1990).
RN [3]
RP SEQUENCE FROM N.A.
RA Dunst R.W., Bell P.A., English J.D.;
RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL: 078873; AAB37349.1; -
DR HSSP: P08515; IGNE.
DR INTERPRO: IPR000521; -
DR PFAM: PF00043; GST; 1.
KW Transferase; Plasmid
SQ SEQUENCE 243 AA; 27908 MW; 0003C434CD76C3A6 CRC64;

Query Match 65.0%; Score 1202.5; DB 2; Length 243;
Best Local Similarity 94.2%; Pred. No. 5.8e-91;
Matches 228; Conservative 1; Mismatches 6; Indels 7; Gaps 2;

OY 1 MSPILGYWKIKGLVQPTRLLEYLEEKYEHLERDEGDKWRKKFKELGLEFPNLPYYID 60
DB 1 MSPILGYWKIKGLVQPTRLLEYLEEKYEHLERDEGDKWRKKFKELGLEFPNLPYYID 60
OY 61 GGVKLTQSMALIRYIADKHNMLGGCCKERAETSMLEGAVIDIRYGSRTAYSKDFETLKV 120
DB 61 GGVKLTQSMALIRYIADKHNMLGGCCKERAETSMLEGAVIDIRYGSRTAYSKDFETLKV 120
OY 121 DFLSKLEPMKMFEDRLCHKTYLNGDVTHTPDMLYDALDVLYNDPMDLAFPKLVCFK 180
DB 121 DFLSKLEPMKMFEDRLCHKTYLNGDVTHTPDMLYDALDVLYNDPMDLAFPKLVCFK 180
OY 181 KRIEALPQIDKYLKSSKYIAMPLOGWQATFGGDDHPKSDLV-----PRGSPNSRVDSG 235
DB 181 KRIEALPQIDKYLKSSKYIAMPLOGWQATFGGDDHPKSDLV-----IPGST 238
OY 236 RA 237
DB 239 RA 240

RESULT 5
CO09191 PRELIMINARY; PRT; 244 AA.
ID 009191;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE GLUTATHIONE S-TRANSFERASE.
GN GST.
OS Escherichia coli.
OG Plasmid pGEX-6P-1.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxId=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88329742; PubMed=3047011;
RA Smith D.B., Johnson K.S.;
RT "Single-step purification of polypeptides expressed in Escherichia
coli as fusions with glutathione S-transferase.";
RL Gene 67:31-40(1988).
RN [2]
RP SEQUENCE OF 221-227 FROM N.A.
RX MEDLINE=90264387; PubMed=2160953;
RA Cordingley M.G., Callahan P.L., Sardana V.V., Garsky V.M.,
RA Colomo R.J.;
RT "Substrate requirements of human rhinovirus 3C protease for peptide
cleavage in vitro.";
RL J. Biol. Chem. 265:9062-9065(1990).
RN [3]
RP SEQUENCE FROM N.A.
RA Dunst R.W., Bell P.A., English J.D.;
RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.

DR EMBL: 078872; AAB37346.1; -
DR HSSP: P08515; IGNE.
DR INTERPRO: IPR000521; -
DR PFAM: PF00043; GST; 1.
KW Transferase; Plasmid
SQ SEQUENCE 244 AA; 28430 MW; 7461B31F17027042 CRC64;

Query Match 65.0%; Score 1202.5; DB 2; Length 244;
Best Local Similarity 93.8%; Pred. No. 5.8e-91;
Matches 228; Conservative 1; Mismatches 5; Indels 9; Gaps 2;

OY 1 MSPILGYWKIKGLVQPTRLLEYLEEKYEHLERDEGDKWRKKFKELGLEFPNLPYYID 60
DB 1 MSPILGYWKIKGLVQPTRLLEYLEEKYEHLERDEGDKWRKKFKELGLEFPNLPYYID 60
OY 61 GGVKLTQSMALIRYIADKHNMLGGCCKERAETSMLEGAVIDIRYGSRTAYSKDFETLKV 120
DB 61 GGVKLTQSMALIRYIADKHNMLGGCCKERAETSMLEGAVIDIRYGSRTAYSKDFETLKV 120
OY 121 DFLSKLEPMKMFEDRLCHKTYLNGDVTHTPDMLYDALDVLYNDPMDLAFPKLVCFK 180
DB 121 DFLSKLEPMKMFEDRLCHKTYLNGDVTHTPDMLYDALDVLYNDPMDLAFPKLVCFK 180
OY 181 KRIEALPQIDKYLKSSKYIAMPLOGWQATFGGDDHPKSDLV-----PRGSPNSRVDSG 235
DB 181 KRIEALPQIDKYLKSSKYIAMPLOGWQATFGGDDHPKSDLV-----EFGP 236
OY 236 RAE 238
DB 237 RLE 239

RESULT 6
Q94745 PRELIMINARY; PRT; 218 AA.
ID Q94745;
AC Q94745;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE 26 KDA GLUTATHIONE S-TRANSFERASE.
OS Schistosoma mekongi.
OC Strigoida; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea;
OC Eukaryota; Metazoa; Platyhelminthes; Schistosomatidae; Schistosoma.
OX NCBI_TaxId=38744;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=THAILAND;
RX MEDLINE=98222395; PubMed=9561610;
RA Grams S.V., Grams R., Korge G., Viyanant V., Upatham S.;
RT "Cloning and sequence analysis of the 26 kDa glutathione S-transferase
gene of Schistosoma mekongi.";
RL Southeast Asian J. Trop. Med. Public Health 28:570-574(1997).
DR EMBL: Y07663; CAA68944.1; -
DR HSSP: P08515; IGTA.
DR INTERPRO: IPR000521; -
DR PFAM: PF00043; GST; 1.
KW Transferase.
SQ SEQUENCE 218 AA; 25486 MW; D3080620B19DE23A CRC64;

Query Match 56.7%; Score 1049; DB 5; Length 218;
Best Local Similarity 89.4%; Pred. No. 1.9e-78;
Matches 195; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

OY 1 MSPILGYWKIKGLVQPTRLLEYLEEKYEHLERDEGDKWRKKFKELGLEFPNLPYYID 60
DB 1 MSPILGYWKIKGLVQPTRLLEYLEEKYEHLERDEGDKWRKKFKELGLEFPNLPYYID 60
OY 61 GGVKLTQSMALIRYIADKHNMLGGCCKERAETSMLEGAVIDIRYGSRTAYSKDFETLKV 120
DB 61 GGVKLTQSMALIRYIADKHNMLGGCCKERAETSMLEGAVIDIRYGSRTAYSKDFETLKV 120

```

OY 121 DFLSKLPKMFEDRLCKHTYLNCDHYTHDPFMDALDVLVYMDPMDLAFPKLVCFK 180
DB 121 DFLNLPKMFEDRLCKHTYLNCDHYTHDPFMDALDVLVYMDPMDLAFPKLVCFK 180
OY 181 KRIEAIPOIDKYLKSSKYIAMPLOGQATFGGDDHPK 218
DB 181 KRIENIPOINEYLKSSKYIEMFLOGQATFGGDDHPK 218

RESULT 7
OY 025595 PRELIMINARY; PRT: 218 AA.
AC 025595;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1999 (TremBLrel. 12, last sequence update)
DT 01-MAY-2000 (TremBLrel. 13, last annotation update)
DE PUTATIVE GLUTATHIONE TRANSFERASE.
OS Clonorchis sinensis.
OC Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea;
OC Opisthorchidae; Opisthorchidae; Opisthorchidae; Opisthorchidae;
OC Clonorchis.
OX NCBI_TaxID=79923;
RN [1]
RP SEQUENCE FROM N.A.
RA Hong S.-J., Lee D.-H.;
RT "Cloning and over expression of 26 kDa glutathione S-transferase from
RL Thesis (1996), Parasitology, Chung-Ang University College of Medicine,
RN Seoul.
RN (2)
RP SEQUENCE FROM N.A.
RA Hong S.-J.;
RT Submitted (APR-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: L47992; AAB46369.3;
DR HSSP: P31670; IPHE.
DR INTERPRO: IPR000521;
DR PFM: PFM00043; GST; 1.
KM Transferase.
SQ SEQUENCE 218 AA; 25038 MW; 7CB17C7B837A0B7C CRC64;

Query Match 38.8%; Score 718; DB 5; Length 218;
Best Local Similarity 59.2%; Pred. No. 2.6e-51;
Matches 129; Conservative 37; Mismatches 52; Indels 0; Gaps 0;

OY 1 MSPILGYWKIKGLVOPTRLLLEYLEKYEHLHYERDEGDKMKKFELEPPNLPYYID 60
DB 1 MAPVIGYWKIRGLADPIRLLELYGDSYEHSHYGRCDGSKMNDKHNGLLEPNLPYYID 60
OY 61 GGVKLTQSMALIRYIADKHNMLGGCPKERAETISMLEGAVLDIRYGVSRIVASKDEFETLKV 120
DB 61 GNFSLTQSLAIRYIADKHNMLGCPKERAETISMLEGAVLDIRYGVSRIVASKDEFETLKV 120
OY 121 DFLSKLPKMFEDRLCKHTYLNCDHYTHDPFMDALDVLVYMDPMDLAFPKLVCFK 180
DB 121 DFLNLPKMFEDRLCKHTYLNCDHYTHDPFMDALDVLVYMDPMDLAFPKLVCFK 180
OY 181 KRIEAIPOIDKYLKSSKYIAMPLOGQATFGGDDHPK 218
DB 181 KRIENIPOINEYLKSSKYIEMFLOGQATFGGDDHPK 218

RESULT 8
OY 09XYL9 PRELIMINARY; PRT: 218 AA.
AC 09XYL9;
DT 01-NOV-1999 (TremBLrel. 12, Created)
DT 01-NOV-1999 (TremBLrel. 12, last sequence update)
DT 01-OCT-2000 (TremBLrel. 15, last annotation update)
DE GLUTATHIONE S-TRANSFERASE.
OS GST-1.
OC Fasciola gigantica.
OC Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea;

```

```

OC Echinostomidae; Echinostomata; Fasciolidae; Fasciolidae; Fasciola.
OX NCBI_TaxID=46835;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-THAILAND.
RA Grams S.V., Grams R., Sobhon P., Vivanant V., Upatham E.S.;
RT "Molecular cloning of expressed antigens from Fasciola gigantica.";
RL Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF112567; AAD23997.1;
DR HSSP: P31670; IPHE.
DR INTERPRO: IPR000521;
DR PFM: PFM00043; GST; 1.
KM Transferase.
SQ SEQUENCE 218 AA; 25337 MW; E9423D75C3F2EAEF CRC64;

Query Match 37.7%; Score 698; DB 5; Length 218;
Best Local Similarity 56.7%; Pred. No. 1.1e-49;
Matches 125; Conservative 35; Mismatches 53; Indels 0; Gaps 0;

OY 5 LGYWKIKGLVOPTRLLLEYLEKYEHLHYERDEGDKMKKFELEPPNLPYYIDGDK 64
DB 5 LGYWKIKGLVOPTRLLLEYLEKYEHLHYERDEGDKMKKFELEPPNLPYYIDGDK 64
OY 65 LTQSMALIRYIADKHNMLGGCPKERAETISMLEGAVLDIRYGVSRIVASKDEFETLKV 124
DB 65 LTQSMALIRYIADKHNMLGGCPKERAETISMLEGAVLDIRYGVSRIVASKDEFETLKV 124
OY 125 KLPEMLKMFEDRLCKHTYLNCDHYTHDPFMDALDVLVYMDPMDLAFPKLVCFK 184
DB 125 ELPKTKMMSDFLGRQYLTGSSVSHVDFMWEALDCTIRYLAPOCLNDFPKLEKRSIE 184
OY 185 AIPQIDKYLKSSKYIAMPLOGQATFGGDDHPK 217
DB 185 DLPIKAYMESKPIKMPPLNLSWTASFSGGDAAP 217

RESULT 9
OY 09NOV4 PRELIMINARY; PRT: 218 AA.
AC 09NOV4;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, last sequence update)
DT 01-OCT-2000 (TremBLrel. 15, last annotation update)
DE CLASS MU GLUTATHIONE S-TRANSFERASE.
OS GSTM.
OC Bos taurus (bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Tissue-eye lens.
RA Jimenez-Jasensio J.V., Garland D.;
RT "A lens glutathione S-transferase, class mu, with thiol-specific
RT antioxidant activity.";
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF249588; AAF64308.1;
KM Transferase.
SQ SEQUENCE 218 AA; 25635 MW; 3D02EAF3C07B0A CRC64;

Query Match 27.7%; Score 512.5; DB 6; Length 218;
Best Local Similarity 47.6%; Pred. No. 1.8e-34;
Matches 99; Conservative 35; Mismatches 69; Indels 5; Gaps 1;

OY 1 MSPILGYWKIKGLVOPTRLLLEYLEKYEHLHYERDEG----DKRNKKFELGEPNLP 55
DB 1 MPMILGYWKIRGLADPIRLLELYGDSYEHSHYGRCDGSKMNDKHNGLLEPNLPYYID 60
OY 56 PYIIGDVKLTQSMALIRYIADKHNMLGGCPKERAETISMLEGAVLDIRYGVSRIVSKDEF 115
DB 56 PYIIGDVKLTQSMALIRYIADKHNMLGGCPKERAETISMLEGAVLDIRYGVSRIVSKDEF 115

```

```
DB 61 PYIDGTHKLTOSNALLRIYARKHNLGCTEEMIRVDILENOVMDRLAMARICYSPPD 120
OY 116 ETLKVDLSKLPMLKMFEDRLCHKTYLNGDHYTHDPFMDLADVLYWDMPCLDAPFK 175
DB 121 EKKRPEFLKEIPKIKLSEFLGKRPWFAGDKLTYWDFLYVDLDMHRIREFPCLDAPFN 180
OY 176 LVCFKRIEAIPOIDKYLKSSKIYAMP 203
DB 181 LKDFISREFGLKRIKISAYMKSSRFLPGL 208

RESULT 10
097117 PRELIMINARY: PRT: 223 AA.
ID 097117
AC 097117
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE GLUTATHIONE S-TRANSFERASE.
OS Boophilus microplus (Cattle tick).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodida; Ixodidae; Boophilus.
OX NCBI_Taxid=6941;
RN [1]
RP SEQUENCE FROM N.A.
RA He H., Chen A.C.;
RT "Characterization and molecular cloning of a glutathione S-transferase
RT from larvae of the cattle tick Boophilus microplus (Acari:
RT Ixodidae)".
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF077609; AAD15991.1; -
DR HSSP: P20136; 1GSU.
DR INTERPRO: IPR000521; -
DR INTERPRO: IPR003081; -
DR PFAM: PF00043; GST.1.
DR PRINTS: PR01267; GSTNSFRASEM.
DR TRANSFERASE.
KW SEQUENCE
SQ 223 AA: 25604 MW; F276ACE2D8259174 CRC64;

Query Match 27.2%; Score 502.5; DB 5; Length 223;
Best Local Similarity 47.2%; Pred. No. 1.2e-33;
Matches 102; Conservative 29; Mismatches 80; Indels 5; Gaps 1;

OY 1 MSPILGYMKIKGLVQPTRLLEYLEEKEYEHLERD-----EGDKMKNKFEGLPEPNL 55
DB 1 MAVYLGWMDIRGLAOPRIKLLAHVDAKVDKRTTCGPRPFDRSSMNLNETKIGLEPNL 60
OY 56 PYIIDGVKLTOSMAIIRYIADKHNMLGCCPKERAISMLEGAVLDIRGVSRIVASKDF 115
DB 61 PYIIDGVKLTOSMAIIRYIADKHNMLGCCPKERAISMLEGAVLDIRGVSRIVASKDF 120
OY 116 ETLKVDLSKLPMLKMFEDRLCHKTYLNGDHYTHDPFMDLADVLYWDMPCLDAPFK 175
DB 121 EKKRPEFLKEIPKIKLSEFLGKRPWFAGDKLTYWDFLYVDLDMHRIREFPCLDAPFN 180
OY 176 LVCFKRIEAIPOIDKYLKSSKIYAMPLOGQWATFG 211
DB 181 LKAFVDRIEALPHVAAYLKSDCKIKPILNDMSFG 216

RESULT 11
035660 PRELIMINARY: PRT: 218 AA.
ID 035660
AC 035660
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE GLUTATHIONE S-TRANSFERASE 6 (EC 2.5.1.18) (GLUTATHIONE-S-TRANSFERASE
DE CLASS M5).
GN GSTM6 OR GSTM5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC STRAIN=FVB, 129/SV, AND C57BL/6; TISSUE=LIVER;
RX MEDLINE=98149729; PubMed=9480867;
RA De Bruin W.C.C., te Morsche R.H.M., Wagenmans M.J.M., Alfertink J.C.,
RA Townsend A.J., Wieringa B., Peters W.H.M.;
RT "Identification of a novel murine glutathione S-transferase class mu
RT gene.";
RL Biochem. J. 330:623-626(1998).
CC -1- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER OF
CC EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES (BY
CC SIMILARITY).
CC -1- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-G.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN LIVER, STOMACH AND SMALL
CC INTESTINE. NOT EXPRESSED IN SPLEEN, KIDNEY, COLON, HEART, MUSCLE,
CC BRAIN OR LUNG.
CC -1- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
DR EMBL: AJ000413; CA04061.1; -
DR EMBL: AJ000412; CA04060.1; -
DR HSSP: P09488; 1GTU.
DR MGD: MGI:1309467; Gstm6.
DR INTERPRO: IPR000521; -
DR INTERPRO: IPR003081; -
DR PFAM: PF00043; GST.1.
DR PRINTS: PR01267; GSTNSFRASEM.
DR TRANSFERASE; Multigene family.
KW SEQUENCE
SQ 218 AA: 25628 MW; 32AB20CC6C3594B CRC64;
```

```
Query Match 25.9%; Score 478.5; DB 11; Length 218;
Best Local Similarity 44.0%; Pred. No. 1.1e-31;
Matches 95; Conservative 35; Mismatches 81; Indels 5; Gaps 1;

OY 1 MSPILGYMKIKGLVQPTRLLEYLEEKEYEHLERD-----DKMKNKFEGLPEPNL 55
DB 1 MPYTLGWMDIRGLAHARLLELEYETCYERRRAMGAPDYDRSQWMLNDKFKLGDPNL 60
OY 56 PYIIDGVKLTOSMAIIRYIADKHNMLGCCPKERAISMLEGAVLDIRGVSRIVASKDF 115
DB 61 PYIIDGVKLTOSMAIIRYIADKHNMLGCCPKERAISMLEGAVLDIRGVSRIVASKDF 120
OY 116 ETLKVDLSKLPMLKMFEDRLCHKTYLNGDHYTHDPFMDLADVLYWDMPCLDAPFK 175
DB 121 EKKRPEFLKEIPKIKLSEFLGKRPWFAGDKLTYWDFLYVDLDMHRIREFPCLDAPFN 180
OY 176 LVCFKRIEAIPOIDKYLKSSKIYAMPLOGQWATFG 211
DB 181 LKDFMAFEGELKRIKISAYMKTSRFLPSPVLYLKQATWG 216

RESULT 12
097SM5 PRELIMINARY: PRT: 218 AA.
ID 097SM5
AC 097SM5
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE GLUTATHIONE S-TRANSFERASE MU-CLASS SUBUNIT M1 (EC 2.5.1.18).
GN GSTM1.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecidae; Macaca.
OX NCBI_Taxid=9541;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang C., Eaton D.L.;
RT "Mu-class GSTs Are Responsible for Aflatoxin B1-8,9-epoxide Activity
```

RT in the Non-human Primate Macaca fascicularis Liver."
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF200709; AAF08539.1; -
 DR HSSP: P09488; 1GTU.
 DR INTERPRO: IPR000521; -
 DR INTERPRO: IPR003081; -
 DR PFAM: PF00043; GST; 1.
 DR PRINTS: PR01267; GSTRNSFRASEM.
 DR Transferrase.
 SQ SEQUENCE 218 AA: 25577 MW: EESB4D80F60C95EB CRC64:

Query Match 25.8%; Score 477.5; DB 6; Length 218;
 Best Local Similarity 43.5%; Pred. No. 1.4e-31;
 Matches 94; Conservative 36; Mismatches 81; Indels 5; Gaps 1;

OY 1 MSPILGYWKIKGLVQPTRLLEYLEEKYEHLERDEG-----DKRNKKFELGLEFPNL 55
 DB 1 MPTLQYWDIRGLAHARILLETDTSSYEKKTYMGDAPDYSQWLNEKFKGLDPNL 60
 OY 56 PYIDGDKLQTSMAIIRYIADKHNMLGGCPKERAISMLEGAVLDIRYGSRIAYSKDF 115
 DB 61 PYLIDGTHKITQSNALIRYIARKHNLCGTEEEKIRVDILENQAMDVSNQARVCSPDF 120
 OY 116 ETLKVDLFLSKLEPMKMFEDRLCHKTYLNGDHYTHPEMLYDALDVLYMDPCIDAFPK 175
 DB 121 EKLKPEYLEGLEPMLMOHFSQFLGKRPWFVGDKITFEVDLFAVDLDRHFEKCIDAFPN 180
 OY 176 LVCFKKRIEALPQIDKYLKSSKYIAMPLQGWATFG 211
 DB 181 LKDFISRFEGLEKISAYMKSSRFLRPVFTTKMAVWG 216

RESULT 13
 ID O9TSM4 PRELIMINARY; PRT: 218 AA.
 AC O9TSM4;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE GLUTATHIONE S-TRANSFERASE MU-CLASS SUBUNIT M2 (EC 2.5.1.18).
 GN GSTM2.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Wang C., Eaton D.L.;
 RT "Mu-class GSTs are responsible for Aflatoxin B1-8,9-epoxide Activity
 in the Non-human Primate Macaca fascicularis Liver."
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF200710; AAF08540.1; -
 DR HSSP: P28161; 1HNA.
 DR INTERPRO: IPR000521; -
 DR INTERPRO: IPR003081; -
 DR PFAM: PF00043; GST; 1.
 DR PRINTS: PR01267; GSTRNSFRASEM.
 DR Transferrase.
 KW SEQUENCE 218 AA: 25708 MW: F2E509C3949F9051 CRC64;

Query Match 25.6%; Score 474.5; DB 6; Length 218;
 Best Local Similarity 44.4%; Pred. No. 2.4e-31;
 Matches 96; Conservative 34; Mismatches 81; Indels 5; Gaps 1;

OY 1 MSPILGYWKIKGLVQPTRLLEYLEEKYEHLERDEG-----DKRNKKFELGLEFPNL 55
 DB 1 MPTLQYWDIRGLAHARILLETDTSSYEKKTYMGDAPDYSQWLNEKFKGLDPNL 60
 OY 56 PYIDGDKLQTSMAIIRYIADKHNMLGGCPKERAISMLEGAVLDIRYGSRIAYSKDF 115

DB 61 PYLIDGTHKITQSNALIRYIARKHNLCGTEEEKIRVDILENQAMDVSNQARVCSPDF 120
 OY 116 ETLKVDLFLSKLEPMKMFEDRLCHKTYLNGDHYTHPEMLYDALDVLYMDPCIDAFPK 175
 DB 121 EKLKPEYLEGLEPMLMOHFSQFLGKRPWFVGDKITFEVDLFAVDLDRHFEKCIDAFPN 180
 OY 176 LVCFKKRIEALPQIDKYLKSSKYIAMPLQGWATFG 211
 DB 181 LKDFISRFEGLEKISAYMKSSRFLRPVFTTKMAVWG 216

RESULT 14
 ID Q27653 PRELIMINARY; PRT: 219 AA.
 AC Q27653;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE GLUTATHIONE TRANSFERASE (EC 2.5.1.18) (GLUTATHIONE S-ALKYLTRANSFERASE)
 DE (GLUTATHIONE S-ARYLTRANSFERASE) (S-(HYDROXYALKYL)GLUTATHIONE LYASE)
 GN EMGST1.
 OS Eschinococcus multilocularis.
 OC Eukaryota; Metazoa; Platyhelminthes; Neodermata; Cestoda; Eucestoda;
 OC Cyclophyllidae; Taeniidae; Eschinococcus.
 OX NCBI_TaxID=6211;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96379220; PubMed=8784771;
 RA Liebau E., Muller V., Lucius R., Walter R.D., Henkle-Dunnen K.;
 RT "Molecular cloning, expression and characterization of a recombinant
 glutathione S-transferase from Eschinococcus multilocularis";
 RL Mol. Biochem. Parasitol. 77:49-56(1996).
 CC "- CATALYTIC ACTIVITY: RX + GLUTATHIONE = HX + R-S-G.
 DR EMBL: X85736; CAA59739.1; -
 DR HSSP: P20136; 1GSU.
 DR INTERPRO: IPR000521; -
 DR PFAM: PF00043; GST; 1.
 KW Transferrase.
 SO SEQUENCE 219 AA: 25470 MW: 74FA820AC56F745B CRC64;

Query Match 25.6%; Score 474.5; DB 5; Length 219;
 Best Local Similarity 42.4%; Pred. No. 2.4e-31;
 Matches 92; Conservative 40; Mismatches 80; Indels 5; Gaps 1;

OY 1 MSPILGYWKIKGLVQPTRLLEYLEEKYEHLERDEG-----RDGDKRNKKFELGLEFPNL 55
 DB 1 MAPTLAYWDIRGLAEQSRLLIKYLEVEYDCKRYKIGSAPFPRDSAMLSSEKFSGLDPNL 60
 OY 56 PYIDGDKLQTSMAIIRYIADKHNMLGGCPKERAISMLEGAVLDIRYGSRIAYSKDF 115
 DB 61 PYIDGDKLQTSMAIIRYIADKHNMLGGCPKERAISMLEGAVLDIRYGSRIAYSKDF 120
 OY 116 ETLKVDLFLSKLEPMKMFEDRLCHKTYLNGDHYTHPEMLYDALDVLYMDPCIDAFPK 175
 DB 121 EKLKPEYLEGLEPMLMOHFSQFLGKRPWFVGDKITFEVDLFAVDLDRHFEKCIDAFPN 180
 OY 176 LVCFKKRIEALPQIDKYLKSSKYIAMPLQGWATFG 212
 DB 181 LKAVLSRFENLPAIRDYMASKEFKTCPCNGASAKMWG 217

RESULT 15
 ID O16058 PRELIMINARY; PRT: 219 AA.
 AC O16058;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE GLUTATHIONE S-TRANSFERASE.
 GN GSTMU.

OS Echinococcus granulosus.
OC Eukaryota; Metazoa; Platyhelminthes; Neodermata; Cestoda; Eucestoda;
OC Cyclophyllidae; Taeniidae; Echinococcus.
OX NCBI_TaxId=6210;
RN (1)
RP SEQUENCE FROM N.A.
RA Fernandez V., Zaha A., Fernandez C.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RA Fernandez Mancebo V., Chalar C., Martinez C., Zaha A.,
RA Fernandez Granja C.;
RT "EGGST: a glutathione S-transferase gene from Echinococcus
granulosus.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF005928; AAB6318.1; -.
DR EMBL; AF101269; AAD16438.1; -.
DR HSSP; P20136; 1GSU.
DR INTERPRO; IPR000521; -.
DR PFM; PF00043; GST; 1.
KW Transferase.
SQ SEQUENCE 219 AA; 25553 MW; 8883E70AD075D154 CRC64;

. Query Match 25.6%; Score 473.5; DB 5; length 219;
Best Local Similarity 42.4%; Pred. No. 2.9e-31;
Matches 92; Conservative 40; Mismatches 80; Indels 5; Gaps 1;
QY 1 MSPIIGYMKIKGVOPTRILLELEKEYEHLVE-----RDEGDKWRNKKFELGLEFPNL 55
Db 1 MAPTLAYWDIRGLAESRLILKYLEVEDDKRYKIGSTPTEDRSAMLSKPSLGLDFPNL 60
QY 56 PYIIDGVKLTOSMAIRYIADKHNMLGCGPKRAEISMLEGAVLDIRYGVSRIRAYSKDF 115
Db 61 PYIIDGDFLTOSGALILEYIADRHGMIPDCKKRAVLIHMLQCEVVDLRMAFTRTCYSPDF 120
QY 116 ETLKVDLFLSKLPBMLKMFEDRLCHKTYLNGDVHTHPDEMLYDALDVVLYMDPKCLDAFPR 175
Db 121 EKLKPLGFETLAOKLPNFETAYLGEKEMLTGDKINYPDFSLCELLNOLMKFEPCTLEKYPR 180
QY 176 LVCFKKRIEAIPOIDKYLLKSKXYIAMPLOGWMTFGG 212
Db 181 LKAYLSRFENLPALRDYMASKEFKTRPCNGASAKMRG 217

Search completed: March 18, 2001, 05:37:15
Job time: 463 sec

